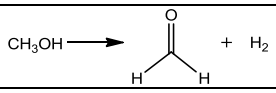
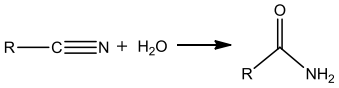
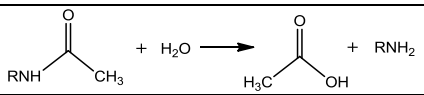
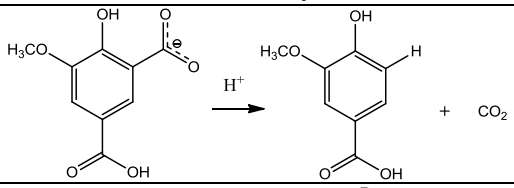
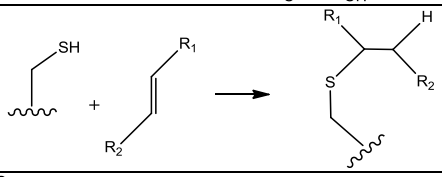
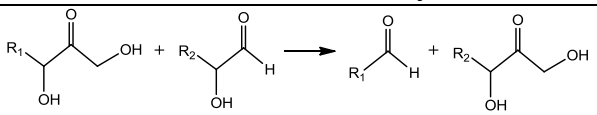
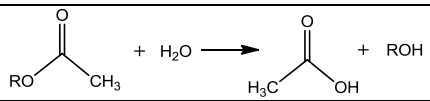
Abstract

Computational Enzymology is became a powerful tools to investigate reactions occurring in enzymatic environment. Nowadays, the modern methodologies coupled to an increasing hardware power allow to investigate several aspects of the interaction between enzyme and substrate, since this last one arrives in proximity of the active site until it is been processed, transformed and released as product. For these reasons, computer simulations are equally relevant to the laboratory experiments, because they provide accurate and detailed aspects at atomistic level. These information are subsequently used in many fields, like design of biomimetics species (biocatalysts) and of new drug (more than 50% of commercial drugs are enzyme inhibitors). The present thesis is based on studies about enzymes belonging to different enzymatic classes and adopting many tools provided by modern theoretical and computational chemistry. Coupling quantum mechanics (QM) based on the Density of Functional Theory (DFT) and hybrid quantum mechanics/molecular mechanics (QM/MM) methods, with Molecular Dynamics (MD) simulations, the study of the reaction mechanisms of the selected enzymes has been deeply addressed and the role of the protein surrounding has been evaluated. In **Chapter 1** a brief outline of theoretical background with the different applied computational approaches is given. In the **Chapter 2** are highlighted relevant aspects of modern enzymology, starting from historical background. The results from the calculations on the specific enzymes are presented in **Chapter 3,** in particular**:**

- In the *session 1* are reported results by QM/MM and/or QM investigations for four metalloenzymes: the methanol-de-hydrogenase (MDH, EC 1.1.2.7), the nitrile hydratase (NHase, EC 4.2.1.84), the peptide glycan N-polysaccharide deacetylase (PDA, EC 3.5.1.104), and 5-carboxyvanillate de-carboxylase (LigW, EC 3.1). In the case of MDH and NHase enzymes, the catalytic mechanisms, the electronic properties and the role of the amino acid residues surrounding the active site have been evaluated depending on the nature of the metal Ca/Ce for MDH (**Paper I)** and Fe/Co–for NHase, (**Paper II**). For the Zn dependent PDA enzyme, deeper insights on its catalytic power, due to the presence in active site of a hydroxyl proline modified after post translational mutation, have been gained in comparison with that of the natural proline (**Paper III**) and finally in the case of the LigW, a Mn dependent enzyme, different models and different approaches have been evaluated with the aim to get information about how theoretical kinetics parameters variate. (**Paper IV**)
- In the *session 2*, the human transketolase (hTK, EC 2.2.1.1), an enzyme presenting the cofactor thiamine diphosphate as co-catalytic agent, has undergone to an accurate study, adopting different methodologies, in order to quantify the ground state de-stabilization effect to the catalysis. (**Work in progress 1**)
- In the *session 3*, QM, QMMM and MD results about glutathione S-transferase Pi 1 (GSTP1, EC 2.5.1.18), as promising enzyme target of the inhibitor piperlongumine (PL), a new potential anti-cancer drug are given (**Paper V**)
- Finally in the *section 4*, the catalytic mechanism of a fully *de novo* protein is presented. Results obtained for this engineered system show as the adopted computational methodologies can provide important information for its possible catalytic use. (**Work in progress 2**)

	Enzyme	EC	Organism	Reaction	Methods
Paper I	Methanol Dehydrogenase	1.1.2.7	*Methylacidiphilum fumariolicum SolV*		Full QM
Paper II	Nitrile hydratase	4.2.1.84	*Rhodococcus erythropolis Pseudonocardia thermophila*		Full QM
Paper III	Peptide glycan N-polisaccaride deacetylase	3.5.1.104	*Bacillus cereus*		MD Full QM
Paper IV	5-carboxyvanillate decarboxylase	3.1	*Novosphingobium aromaticivorans*		Full QM QMMM
Paper V	Glutathione S-transferase Pi 1	2.5.1.18	*Homo Sapiens*		MD Full QM QMMM
Wp 1	Human tranketolase	2.2.1.1	*Homo sapiens*		MD Full QM QMMM
Wp 2	CC-Hept-Cys$_{18}$-His$_{22}$-Glu$_{25}$	-	-		MD QMMM

Contents

Acknowledgements **5**
List of publications **6**
Amino Acid Abbreviations **8**
List of Abbreviations and Acronyms **9**
Chapter 1 ..... **10**
1.1 Introduction 10
1.2 Theory 11
1.2.1 Density Functional Theory 11
1.2.2 Molecular Dynamics 14
1.2.3 Molecular Docking 15
1.3 Methods 16
1.3.1 Full QM methodologies: the Cluster Approach 16
1.3.2 Hybrid QM/MM and ONIOM methodologies 17
References 19
Chapter 2 ..... **22**
2.1 Introduction 22
2.2 One hundred years of models 23
2.2.1 Michaelis Menten Kinetic 24
2.3 Cofactors 25
2.4 Catalysis 26
2.5 Inhibition 27
2.6 *De Novo* proteins 28
References 29
Chapter 3 ..... **32**
Section 1 32
3.1 Metalloenzymes 32
3.1.1 Cerium Methanol de-hydrogenase (Paper I) 32
3.1.2 Nitrile Hydratase containing Iron and Cobalt (Paper II) 34
3.1.3 Peptidoglycan N-acetylglucosamine deacetylase containing Zinc (Paper III) 37
3.1.4 5-carboxyvanillate decarboxylase containing Manganese (Paper IV) 37
Section 2 40
3.2 human Transketolase (Wp 1) 40
Section 3 40
3.3 Inhibition of glutathione S-transferase (Paper V) 40
Section 4 40
3.4 *De novo* protein (Wp 2) 40
References 41

Concluding Remarks **43**
Paper I ..... **A**
Paper II ..... **B**
Paper III ..... **C**
Paper IV ..... **D**
Paper V ..... **E**
Wp 1 ..... **F**
Wp 2 ..... **G**

Acknowledgements

First of all, I express my gratitude to the past and present members of PROMOCS laboratory at Department of Chemistry and Chemical Technologies of University of Calabria, which supported me with their friendly and collaborative presence, in the last three years.

My sincere thanks to my supervisor Prof. Tiziana Marino, for her patience and strong will to share with me her knowledge, for all those stimulating discussions and opportunities for personal growth.

I would thank Prof. Nino Russo for his guidance. His encouragement was continued spur to keep going with scientific curiosity and to hard work.

Thanks to Prof. Maria João Ramos, who gave me the opportunity to spend a research period at the University of Porto, and to Prof. Pedro Alexandrino Fernandes. In particular, thank you for all the inspiring discussions. Thanks to Fabiola Medina and to all the members of Computational Chemistry and Biochemistry Group for their support during my stage in Porto.

List of publications

The results in the present thesis are based on following papers.

I. **How Can Methanol Dehydrogenase from Methylacidiphilum fumariolicum Work with the Alien Ce^{III} Ion in the Active Center? A Theoretical Study**
Mario Prejanò, Tiziana Marino, Nino Russo
Chem. Eur. J. **2017**, *23*, 8652-8657.
doi:10.1002/chem.201700381

II. **Reaction Mechanism of Low-Spin Iron(III)- and Cobalt(III)-Containing Nitrile Hydratases: A Quantum Mechanics Investigation**
Mario Prejanò, Tiziana Marino, Carmen Rizzuto, Josè Carlos Madrid Madrid, Nino Russo, Marirosa Toscano.
Inorg. Chem. **2017**, *56,* 13390-13400.
doi: 10.1021/acs.inorgchem.7b02121

III. **Why nature prefers hydroxy-proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations**
Luigi Sgrizzi, Mario Prejanò, Isabella Romeo, Tiziana Marino, Nino Russo
Manuscript submitted

IV. **QM Cluster or QM/MM in Computational Enzymology: The Test Case of LigW-Decarboxylase**
Mario Prejanò, Tiziana Marino, Nino Russo
Frontiers in chemistry, **2018**, *6*, 249-257.
doi: 10.3389/fchem.2018.00249

V. **On the inhibition mechanism of glutathione transferase P1 by piperlongumine. Insight from theory**
Mario Prejanò, Tiziana Marino, Nino Russo
Manuscript submitted and accepted

Work in progress (Wp)

Wp 1. **How the destabilization of a reaction intermediate affects enzymatic catalysis: the case of human transketolase**
Mario Prejanò, Fabiola E. Medina, Tiziana Marino, Pedro Alexandrino Fernandes, Maria Joao Ramos, Nino Russo
Manuscript in preparation

Wp 2. **Mechanistic insights of hydrolytic activity into a de novo Functional protein framework.**
Mario Prejanò, Isabella Romeo, Tiziana Marino, Nino Russo
Work in progress

The contribution by the author to the papers is summarized as follows:

Paper I Performed the calculations, analyzed the results and took part in the writing of the article.

Paper II Performed a part of the calculations, analyzed the results and took part in the writing of the article.

Paper III Performed the QM calculations, analyzed the results and took part in the writing of the article.

Paper IV Performed the calculations, analyzed the results and took part in the writing of the article.

Paper V Performed the calculations, analyzed the results and took part in the writing of the article.

Wp 1 Performed the calculations, analyzed the results and took part in the writing of the article.

Wp 2 Performed the QMMM calculations, analyzed the results and took part in the writing of the article.

Amino Acid Abbreviations

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
Cys	Cysteine
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

List of Abbreviations and Acronyms

2-Hyp	2-Hydroxy-Proline
5-CV	5-CarboxyVanillate
AHS	Amidohydrolase
AMBER	Assisted Model Building with Energy Refinement
ATP	Adenosine-Three Phosphate
B3LYP	Becke 3 parameter Lee-Yang-Parr
B88	Becke 88
B95	Becke 95
B97D	Becke 97 Dispersion
CC-Hept	Coiled Coil Heptamer
CC-Hept	Coiled Coil Heptamer
CHARMM	Chemistry at HARvard Molecular Mechanics
CVFF	Consistent Valence Force Field
DFT	Density Functional Theory
E	Enzyme
E4P	Erytrose 4 Phosphate
EC	Enzyme Commission
EI	Enzyme-Inhibitor
EP	Enzyme-Product
ES	Enzyme-Substrate
ESI	Enzyme-Substrate-Inhibitor
F6P	Fructose 6 Phosphate
FF	Force Field
G3P	Glyceraldehyde 3 Phosphate
GlcNAc	Peptidoglycan N-acetylglucosamine
GROMOS	GROningen MOlecular Simulation
GSH	Glutathione
GSTP1	Gluatahione S-Transferase Pi 1
HF	Hartree Fock
HOMO	Highest Occupied Molecular Orbital
hPL	Hydrolyzed PiperLongumine
hTK	Human TransKetolase
KS	Kohn Sham
LUMO	Lowest Unoccupied Molecular Orbital
LYP	Lee Yang Parr
M06	Minnesota 06
M06-2X	Minnesota 06 2 eXchange
M06-L	Minnesota 06 Local
MD	Molecular Dynamics
MDH	Methanol De Hydrogenase
MM	Molecular Mechanics
MM2	Molecular Mechanics 2
mPW	Meta Perdew Wang
mPWB1K	Meta Perdew Wang Becke 95 1
mPWB95	Meta Perdew Wang Becke 95
NAD	Nicotinamide adenine dinucleotide
NBO	Natural Bond Orbital
NHase	Nitrile Hydratase
ONIOM	Our own N-layered Integrated MO and MM
OPLS	Optimized Potential for Liquid Simulations
P	Perdew
P86	Perdew 86
PBE	Perdew Burke Ernzerhof
PBEh	Perdew Burke Ernzerhof hybrid
PDA	Polysaccharide DeAcetylase
PDB	Protein Data Bank
PES	Potential Energy Surfaces
PL	PiperLongumine
pNPA	para Nitro Phenol Acetate
PQQ	Pyrrol Quinoline Quinone
PTM	Post Translational Modification
PW91	Perdew Wang 91
QCFF/PI	Quantum Consistent Force Field PI electrons
QM	Quantum Mechanics
QM/MM	Quantum Mechanics / Molecular Mechanics
REM	Rare Earth Metal
RNA	RiboNucleic Acid
S	Substrate
ThDP	Thiamina Diphosphate
TK	TranKetolase
TS	Transition State
TST	Transition State Theory
UFF	Universal Force Field
VWN	Vosko Wilk Nusair

Chapter 1

Theories and methods

Enzymes, being macro-molecules, are too large to be studied in their totality, applying pure quantum mechanics. According with the type of investigated phenomenon, different theories and different approaches can be used to study enzymes and reactions catalyzed by them. This chapter gives a brief overview of the different theory levels and methods used for the calculations presented in this thesis. In detail, will be reported a summary about density functional theory, molecular docking, molecular mechanics and molecular dynamics, applied to computational enzymology. Coupling full Quantum Mechanics (QM) and hybrid Quantum Mechanics/Molecular Mechanics (QMMM) methods, accurate Potential Energy Surfaces (PESs) to evaluate and predict mechanistic aspects of catalysis and enzyme-inhibitor reactions at atomistic level can be obtained.

1.1 Introduction

The modelling of *reactions in enzymatic environment* may be considered as an important piece of the *Computational Enzymology* puzzle, which, in the same way, is a piece of the *multi-scale* modelling puzzle. This means that, according with the investigated information, and its relative observation-time, specific levels of theory and models may be applied (see **Figure 1**). To study enzymes, the adopted models are different according with the circumstances about the

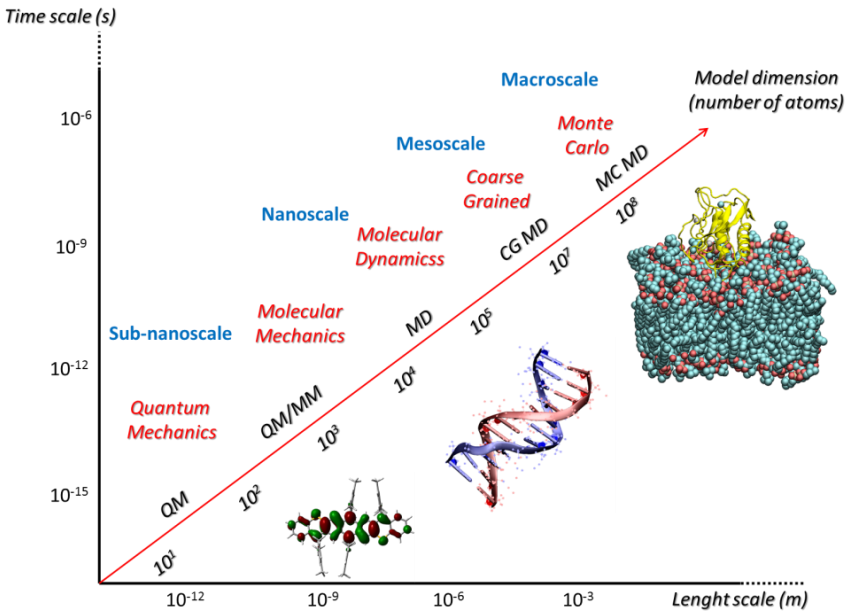
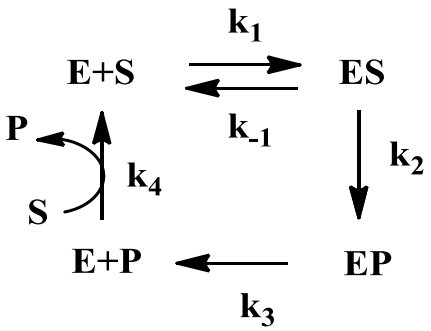


Figure 1. Multi-scale modelling scheme.

interactions between the enzymes and molecules (substrates and inhibitors) involved in reaction, as showed in **Scheme 1**. In the first part, substrate (S) binds reversibly to the enzyme (E) forming the ES complex. In the second step, chemical reactions converting substrate in product (P) happen. Passing from interacting enzyme-product step (EP) to the step without interaction (E+ P), the cycle is restored with the arrival of a new substrate (S). For each step, is possible to build up models with different dimensions and to extrapolate different information.



Scheme 1. Generic catalytic cycle favored by the presence of an enzyme.

In this thesis, attention is focused on the $ES \xrightarrow{k_2} EP$ step, applying. Since the most important aspect of enzymatic reactions is the increase of reaction rate, using the transition state theory (TST) it is possible to theoretically determine energy barrier that in turn can be compared with the experimental one, according to the equation:

$$k_{cat} = \left(\frac{k_B T}{h} \right) e^{\frac{-\Delta G^{**}}{RT}} \qquad \text{(Eq. 1.1)}$$

T is the temperature, k_b and h are the Boltzmann and Planck constants, respectively, R is the universal gas constant and ΔG^{**} is the activation energy for the reaction. ΔG^{**} corresponds to the highest barrier calculated, obtained as a difference between transition state energy and its previous intermediate, once defined the reaction pathway.

1.2 Theory

1.2.1 Density Functional Theory

The quantum mechanics treatment of a chemical system provides the solution of Schrödinger equation. The equation not dependent by the time, in the non-relativistic form, is:

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle \qquad \text{(Eq. 1.2)}$$

For a molecular system with N electrons and M nuclei, the Hamiltonian operator " $\hat{H}$ " is written, in atomic units:

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1,j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1,B>A}^{M}\frac{Z_A Z_B}{r_{AB}} \qquad \text{(Eq. 1.3)}$$

The first and the second terms of equation represent the kinetic energy contribution of electrons and nuclei; the third, fourth and fifth are the potential energy contributions obtained considering electron-nucleus attractive interactions, electron-electron and nucleus-nucleus repulsive interactions, respectively. According with the Born-Oppheneimer[1] approximation is possible to consider separately the nuclei and electrons contributions. Despite the rigorous mathematical formalism, all *ab initio* methods need approximations to solve Schrödinger equation. Several *ab initio* methodologies could be applied, consistently with the system dimension, but the Density Functional Theory (DFT) is still widely used, from the last 20 years to our days, to study complex poly-electrons systems, as the catalytic site of enzymes, thanks to the electron density $\rho(r)$ definition, where *N* is the number of electrons:

$$N = \int \rho(r)dr \qquad \text{(Eq. 1.4)}$$

The modern DFT methods base on the two Hohenberg and Kohn (HK) theorems[2] and the Kohn-Sham (KS) formalism.[3] The first HK theorem demonstrates that the only fundamental variable that determines the Hamiltonian, the eigenvalues correlated to its and the ground-state wave function Ψ, is the electron density $\rho(r)$. The second HK theorem demonstrates that exists a variational principle such that any trial $\rho'(r)$ will result in an energy higher than or equal to the exact ground-state energy. Applying KS formalism is possible to define $\rho(r)$ considering the atomic orbital set ϕ_i:

$$\rho(r) = \sum_{i=1}^{N}\left|\phi_i(r)\right|^2 \qquad \text{(Eq. 1.5)}$$

The electronic energy is expressed as the sum of the three different electron terms, as previously cited:

$$E[\rho(r)] = T_i[\rho(r)] + V_{ij}[\rho(r)] + V_{iA}[\rho(r)] \qquad \text{(Eq. 1.6)}$$

The $V_{ij}[\rho(r)]$ term contains the classical Coulomb integral ($J[\rho(r)]$) and all non-classical electron-electron interactions. $T_i[\rho(r)]$ contains two different kinetic contributions: the first obtained considering a non-interacting electrons system ($T_{ni}[\rho(r)]$) and the second evaluated by the

difference between non-interacting and real electron system. Grouping the last cited energy contribution and the non-classical electron-electron interactions $E[\rho(r)]$ is defined as:

$$E[\rho(r)] = T_{ni}[\rho(r)] + J[\rho(r)] + V_{iA}[\rho(r)] + E_{xc}[\rho(r)] \quad \text{(Eq. 1.7)}$$

Unfortunately, does not exist an exact expression of the exchange-correlation functional $E_{xc}[\rho(r)]$ and various DFT methods have been developed in order to find accurate expression of this functional. DFT functionals can be divided in two main groups, pure and hybrid functionals, according with their obtainment and mathematical expression. Considering $E_{xc}[\rho(r)]$ as follow:

$$E_{xc}[\rho(r)] = E_x[\rho(r)] + E_c[\rho(r)] \quad \text{(Eq. 1.8)}$$

several exchange, correlation and exchange-correlation functionals were proposed. Example are:

- *Correlation functionals:* LYP,[4] B95,[5] PBE,[6] P86 [7] and PW91.[8a]
- *Exchange functionals:* B88,[9] PBE and PBEh,[10] mPW[11] and wPBEh. [12]
- *Pure exchange-correlation functionals:* M06-L[13] and B97D.[14]

About hybrid functionals, the most famous and used are B3LYP,[4,9,15] M06, [16] PBEPBE[17] and wB97XD.[18] Another class of DFT methods is the meta-hybrid functional, with MPWB1K and MPWB95,[19] particularly recommended for thermochemical kinetics studies. [19] As far as the study of enzyme reaction mechanisms by applying the cluster approach, the hybrid functional B3LYP has been extensively applied as it has been considered to be a good trade-off between accuracy and speed. More detailing, the mathematical expression of that hybrid functional is obtained combining the exchange functional Becke[9] (B88) with the gradient corrected correlation of Lee, Yang and Parr (LYP),[4] added to a local-density approximation (VWN)[8b] and 20% of exact HF exchange:

$$E^{B3LYP} = (1-a)E_X^{LSDA} + aE_X^{HF} + bE_X^{B88} + cE_C^{LYP} + (1-c)E_C^{VWN} \quad \text{(Eq. 1.9)}$$

In order to resolve the known weakness of this functional, present also in other functionals, due to their lack of the description of accurately long-range dispersion interactions, since they have local dependency on the electron density, several methods[20] have been developed to include dispersion interaction effect into density functional theory, according to the equation:

$$E_{DFT-D} = E_{DFT} + E_{disp} \quad \text{(Eq. 1.10)}$$

This technique has been shown to yield a significant improvement of the energies in the field of homogeneous catalysis.[21] As well as in enzyme modeling improving both energies and geometries[22] and quickly becoming a standard choice in the cluster approach. Another manner to take into account the dispersion is to use functionals such as the Minnesota M06 suite,[23] that contain weak interactions in the training set. These methods common in the homogeneous catalysis are starting to be used also in enzyme modeling.[24]

1.2.2 Molecular Dynamics

Molecular Dynamics (MD) simulations represent the method of choice when one wants to study the dynamical properties of a system in full atomic detail, provided that the properties are observable within the time scale accessible to simulations. To calculate the dynamics of the system, that is the position of each atom as a function of time, Newton's classical equation of motion are solved iteratively for each atom:

$$F_i = m_i a_i = m_i \frac{d^2 r_i}{dt^2} \quad \text{(Eq. 1.11)}$$

The force on each atom is the negative of the derivative of the potential energy with respect to the position of the atom:

$$F_i = -\frac{\partial V}{\partial r_i} \quad \text{(Eq. 1.12)}$$

If the potential energy of the system is known then, given the coordinates of a starting structure and a set of velocities, the force acting on each atom can be calculated and a new set of coordinates generated, from which new forces can be calculated. Repetition of the procedure will generate a trajectory corresponding to the evolution of the system in time. The accuracy of the simulations is directly related to the potential energy function used to describe the interactions between particles. In molecular dynamics, a classical potential energy function is used that is defined as a function of the coordinates of each of the atoms. The potential energy function is separated into terms representing covalent interactions and non-covalent interactions. The covalent interactions may be described by the following terms:

$$V_{bond} = \sum_{i=1}^{Nb} \frac{k_{bi}}{2} \left(r_i - r_{i,0}\right)^2 \quad \text{(Eq. 1.13)}$$

$$V_{angle} = \sum_{i=1}^{N\theta i} \frac{k_{\theta i}}{2} \left(\theta_i - \theta_{i,0}\right)^2 \quad \text{(Eq. 1.14)}$$

$$V_{dihedral} = \sum_{i=1}^{N\varphi} \frac{k_{\varphi i}}{2} cos\left[n_i\left(\varphi_i - \varphi_{i,0}\right)\right] \quad \text{(Eq. 1.15)}$$

$$V_{improper} = \sum_{i=1}^{N\xi} \frac{k_{\xi i}}{2} \left(\xi_i - \xi_{i,0}\right)^2 \quad \text{(Eq. 1.16)}$$

which correspond to two, three, four and four body interactions, respectively. These interactions are represented by harmonic potentials for the bond lengths r_i , for the bond angle θ_i, and for the

improper dihedral (out of the plane) angle ξ_i and by a more complex potential for the dihedral angles φ_i. The non-covalent interactions, which correspond to interactions between particles separated by more than three covalent bonds are usually described by Coulomb's law

$$V_{Coulomb} = \sum_i \sum_j \frac{q_i q_j}{r_{ij}} \quad \text{(Eq. 1.17)}$$

for the electrostatic interactions and by a Lennard-Jones (V_{LJ}) potential

$$V_{LJ} = \sum_i \sum_j \frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^{6}} \quad \text{(Eq. 1.18)}$$

for the Van der Waals interactions where r_{ij} is the atomic distance between particle i and j. The complete set of parameters used in the potentials (force constants, ideal bond lengths, bond angles, improper dihedral angles, dihedral angles, partial charges and Van der Waals parameters) to describe the interactions between different particle types is called force field (FF). Molecular dynamics is a very useful tool. It can provide a wealth of detailed information on the structure and dynamics of proteins and peptides. These reasons make MD an important preliminary, and/or advanced step, because is fundamental to have optimum starting point in higher level of theory investigation, as will be mentioned successively. Molecular dynamics simulations can provide are also useful information generally not always accessible to the experimental (x-ray or NMR) methods. However, it suffers certain limitations.[25] First, the method is computationally very demanding and depending on the size of the system simulation times are currently limited to hundreds of nanoseconds or a few microseconds at most. Also, the form of the potential energy function must be kept simple for reasons of efficiency. The possibility to observe certain properties is directly related to the quality of the force field and, whether or not it has been parameterized for the system simulated. The quality of the FF is especially critical in the simulation of proteins. The FF thus must be very accurate to discriminate between different conformations. Examples of classical FFs are AMBER, [26] GROMOS,[27] CHARMM,[28] MM2,[29] one of the first proposed, UFF,[30] CVFF,[31] OPLS',[32] and QCFF/PI.[33] However, it is questionable whether an empirical force-field can achieve the required accuracy especially when important effects such as polarization of the atoms by their environment is not taken into account by the electrostatic potential. The last limitation is that a classical description of the particles is used.

1.2.3 Molecular Docking

Molecular Docking simulations have been became widely adopted tools in the last years. When the structure of the enzyme is available in the substrate-free form, molecular docking can be used to dock, or fit, the substrate in the active site.[34,35] The docking algorithm has the ability to predict the conformation of a ligand adopted within the active site of its target. This algorithm has been implemented significantly over the years addressing the flexibility limit of ligand and protein. However, the change in protein conformation induced by external conditions, changes in

compliance with the interaction with the ligand-protein are still unsolved problems. For each analyzed ligand, this algorithm will generate a *score*, calculated empirically or semiempirically, which will try to estimate the complex ligand-protein interactions. This allows to have *a priori* evaluation of interactions involving in an active site, in the case of ES complex, for example, or between two proteins,[36] restituting at the end of calculations relative affinity energies. At the end of simulations, many poses are obtained, considering torsional degrees of freedom of the ligand, in a receptor region adequately chosen. In the present thesis, the followed strategy has been choosing the best calculated conformation by the scoring function and adopting it as starting point for mechanistic studies performed at highest levels of theory.

1.3 Methods

1.3.1 Full QM methodologies: the Cluster Approach

In the *Cluster Approach*, active site models are obtained on the basis of crystal structures made available, for example, in the protein data bank (PDB). Depending from the type of catalytic center, accurate cluster models are used, ranging from 100-150 to 200-300 atoms, with and without metal ions, respectively. The general idea of this full QM investigation is taking into account in the model only a small part of the enzyme to study the properties and reaction mechanism. In the study of an enzymatic reaction, some structural information about the active site represents a necessary prerequisite. Enzymatic reactions are often catalyzed by a number of residues and it is important to know their positions related to each other and to the substrate (if this information is present). For all the systems investigated in this thesis a X-ray structure of the enzyme was available. There are two simple approximations to consider the effects of the rest of the enzyme surrounding: to simulate the whole protein around the catalytic pocket with an implicit solvation model and the, so-called, *coordinate-locking* scheme, where a number of atoms are fixed to their crystallographic positions, typically where truncation is made, i.e. the carbon atom and one or two of the hydrogens that replace the connecting atoms. [37] The use of implicit solvation models to account for the electrostatic influence of the enzyme surrounding assumes that the enzyme surrounding is a homogeneous polarizable medium, with a dielectric constant ε set to equal 4 in most of the examined cases of this thesis. In **Figure 2** are depicted all the steps characterizing the "manipulation" of the X-ray starting structure for generating the QM cluster ready for the QM mechanistic investigation.
Being the case there is no official criteria for selection of amino acid residues to be retained in the QM model, some crucial steps can be resumed:

- if in the active site are present metal ions, their complete coordination sphere must be included in the model;
- amino acids and water molecules believed important to the catalysis, from experimental and/or accurate visual evaluations, for example involved in hydrogen bonds with substrates, must be considered;

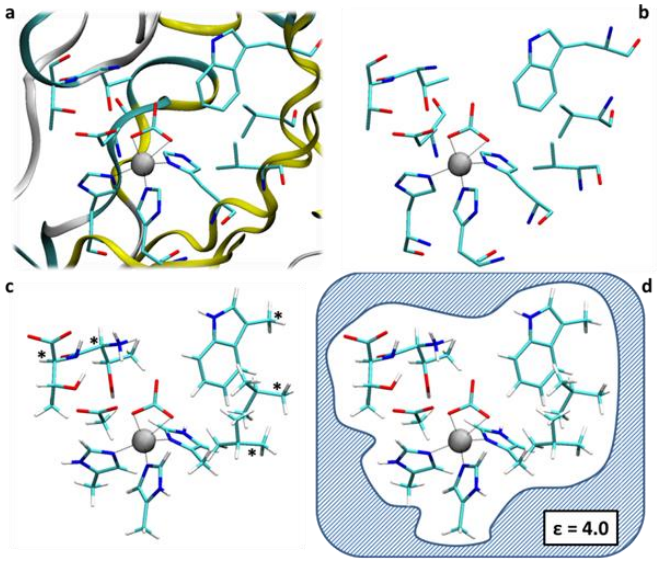


Figure 2. The cluster approach: (a) the active site is modelled starting from X-ray structure available in data banks, selecting amino acid residues important to catalysis (b). The residues are truncated retaining the side chains (c) and the C atoms, marked with "*" are kept frozen during the optimizations. (d) The rest of the enzyme is treated as an homogenous polarizable medium with a dielectric constant.

1.3.2 Hybrid QM/MM and ONIOM methodologies

Inside the QM cluster methodology, neglecting the long range interactions and the role of the protein conformational changes during the reaction could represent a "weakness", in particular if the QM selected model doesn't have appropriate dimensions. A natural way out of this problem was to treat the rest of the enzyme using a simpler, normally classical, description, which was done in the hybrid QM/MM method.[38]

Hybrid QM/MM methodologies differ, mainly, for the following features: [39]

- type of scheme (additive or subtractive);
- how are considered interactions between QM and MM regions;
- the treatment of boundary region.

General equations of the used schemes are additive (eq.19) and subtractive (Eq. 20):

$$E_{QM/MM,totalsystem} = E_{QM.innerlayer} + E_{MM,outlayer} + E_{QM/MM,coupling} \quad \text{(Eq. 1.19)}$$

$$E_{QM/MM,completesystem} = E_{QM,innerlayer} + E_{MM,completesystem} - E_{MM,innerlayer} \quad \text{(Eq. 1.20)}$$

The main difference, in the two approaches, is about the treatment of the part across the QM and MM region, as shown in **Figure 3**. Additive QM/MM methods calculate explicitly the energy of the system as the sum of QM, MM and QM-MM regions.

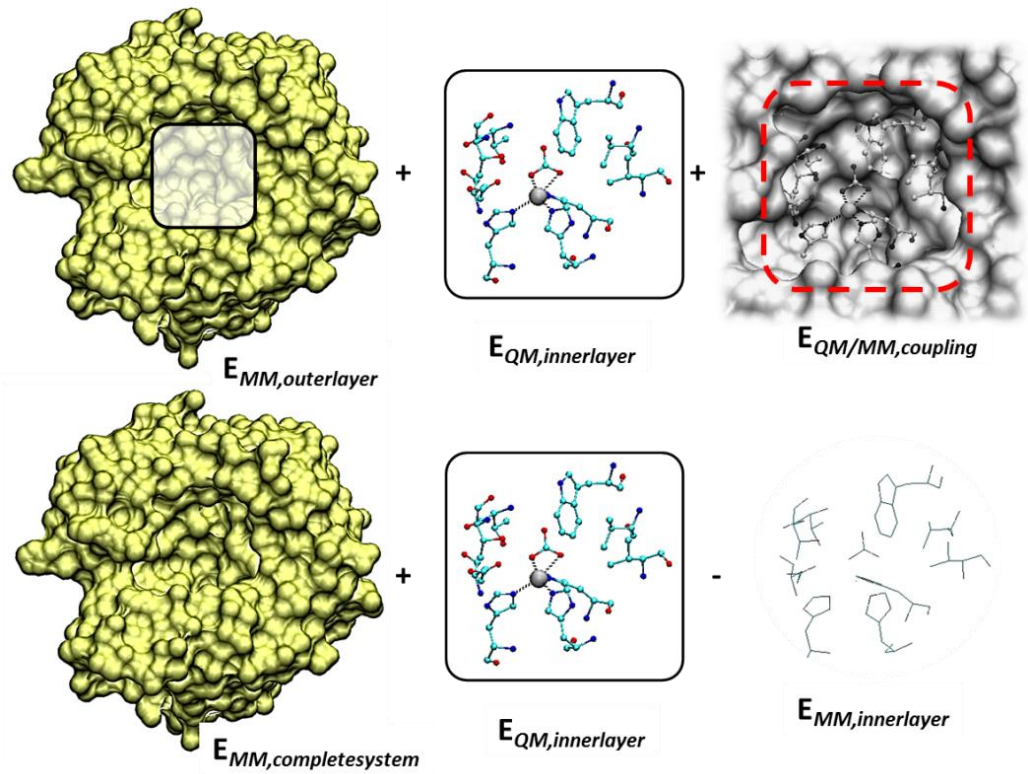


Figure 3. Schematic representation of an additive and a subtractive coupling within a QM/MM methodology. Two-layer additive (top) and subtractive (bottom) schemes.

In first principle, is the most accurate approach and more popular in biomolecular systems, but is difficult to calculate accurately the coupling term $E_{QM/MM\ coupling}$. Generally, in subtractive schemes, QM and MM regions are considered as independent parts of the whole system; in this way, the problem about the accuracy of the coupling region expression is avoided. Morokuma and coworkers have developed three methods[40,41,42] based on a subtractive scheme and, in particular, the "*our own n-layered integrated MO and MM*" method (ONIOM)[42] is that applied in the examined enzymes of the present thesis (**Paper IV**, **Paper V, Wp 1, Wp 2**). In any way other advantages and disadvantages may be encountered performing ONIOM simulations [39,43-46] but they are outside from the purpose of this thesis work. In enzyme QM/MM calculations, there are always covalent bond between QM and MM regions. In this case the two systems cannot betreated directly as separated in the QM/MM calculations. Two common strategies are typically applied to deal with this problem: interaction *link atom* and *frozen atom approach.* In the first case, the valence of the QM atoms covalently bonded to the MM region is filled to its classical coordination number with 'link atoms.' These 'link' atoms are placed to represent the behavior of the MM atoms that are substituted. Commonly in this approach hydrogen atoms are chosen. [47-49] The 'link' atoms are treated as real atoms and explicitly threated in the QM calculations, while in the MM calculation they are treated as a pseudoatom. In the *frozen orbital approach*, a frozen localized orbital is used to describe interface between the two regions, assuming that the bonds have constant and well-defined properties. The bond is obtained from a linear combination between hybrid orbitals of the two atoms. A critical point in QM/MM simulations is handly the electrostatic interactions between the

two regions. Three methodologies are usually adopted: *mechanical, electrostatic or polarizable embedding*. Essentially, *mechanical embedding* does not consider interactions between the two parts; adopting this strategy, accurate MM parameters are requested for QM regions, stationary point by stationary point on the PES, being the atom-centered point charges changing during the reaction. Therefore, to use a fixed set of point charges may generate critical discontinuities in the PESs. With *electrostatic embedding*, polarization of the QM region by the MM charge distribution is considered during QM energy calculation, including one electron in QM Hamiltonian. The inner region is automatically polarized detecting the effect of the charge distribution of the environment and the MM environment is influenced by the new charges of QM part. This method does not consider the polarization of MM system that, in some cases, can have effect in the results. The last methodology named the *polarizable embedding*, finally includes the polarization of the MM region as response to the QM charge distribution. Anyway, being the most accurate way to consider the electrostatic interaction, as it to react to a perturbation by the QM region, a feature that is not typically available in current commercial MM FFs, even though some efforts have been reported.[50-53]

References

1. Born, M.; Oppenheimer, J. R. *Annalen der Physik* **1927**, *84*, 457.
2. Hohenberg, P.; Kohn,W. *Phys. Rev.* **1964**, *136*, B864.
3. Kohn, W.; Sham, L. *J. Phys. Rev.* **1965**, *140*, A1133.
4. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
5. Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.
6. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
7. Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
8. (a) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671; (b) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
9. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
10. a) Ernzerhof, M.; Perdew, J. P. *J. Chem. Phys.* **1998**, *109*, 3313. b) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
11. Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.
12. Heyd, J.; Scuseria, G.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207.
13. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 1.
14. Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787.
15. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
16. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
17. Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
18. Chai, J.-D., Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
19. Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A*, **2004**, *108*, 6908.
20. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
21. a) Minenkov, Y.; Occhipinti, G.; Jensen, V. R. *J. Phys. Chem. A* **2009**, *113*, 11833. b) Harvey, J. N. *Faraday Discuss.* **2010**, *145*, 487. c) McMullin, C. L.; Jover, J.; Harvey, J. N.;

Fey, N. *Dalton Trans* **2010**, *39*, 10833. d) Osuna, S.; Swart, M.; Sola, M. *J. Phys. Chem. A* **2011**, *115*, 3491. e) Santoro, S.; Liao, R.- Z.; Himo, F. *J. Org. Chem.* **2011**, *76*, 9246.
22. a) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3232. b) Siegbahn, P. E. M.; Blomberg, M. R. A.; Chen, S.-L. *J. Chem. Theory Comput.* **2010**, *6*, 2040. c) Zhang, H.-M.; Chen, S.-L. *J. Chem. Theory Comput.* **2015**, *11*, 2525.
23. Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.
24. van Severen, M.-C.; Andrejic, M.; Li, J.; Starke, K.; Mata, R. A.; Nordlander, E.; Ryde, U. *JBIC, J. Biol. Inorg. Chem.* **2014**, *19*, 1165.
25. Karplus, M.; McCammon, A. *Nature structural biology* **2002**, *9*, 646.
26. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. Jr; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
27. Oostenbrink, C; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comp. Chem*. **2004**, *25*, 1656.
28. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
29. Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.
30. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
31. Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe,, D. J.; Wolff, J.; Genest., M.; Hagler, A. T. *Prot.: Struct. Func. And Gen.* **1988**, *4*, 31.
32. Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.***1988**, *110*, 1657.
33. Warshel, A. *Israel J. Chem.* **1973**, *11*, 709.
34. Lengauer, T.; Rarey, M. *Curr. Opin. Struct. Biol.* **1996**, *6*, 402.
35. Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D.; Brooks, C. *J. Comp. Chem*. **2004**, *25*, 265.
36. Hospital, H.; Goñi, J. R.; Orozco, M.; Gelpi, J. L. *Advances and Applications in Bioinformatics and Chemistry* **2015**, *8*, 37.
37. Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R. Z.; Siegbahn, P. E. M. *Chem. Rev.* **2014**, *114*, 3601.
38. Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 139.
39. a) Sousa, S. F. M.; Ribeiro, A. J.; Neves, R. R. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; M. J. Ramos, M. J. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1281. b) Ryde, U. *Methods Enzymol.* **2016** 577, 119; (c) Quesne, M. G.; Borowski, T.; de Visser, S. P. *Chem. Eur. J.* **2016**, *22*, 2562.
40. Maseras, F.; Morokuma. *J. Comput. Chem.* **1995**, *16*, 1170.
41. Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys*. **1996**, *105*, 1959.
42. Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
43. Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
44. Ryde, U. *J. Chem. Theory Comput.* **2017**,*13*, 5745.
45. van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708.
46. Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H-B; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, *115*, 5678.
47. Antes, I.; Thiel, W. *J. Phys. Chem.* **1999**, *103*, 9290.

48. Zhang, Y. K.; Lee, T. S.; Yang, W. T. *J. Phys. Chem*, **1999**, *110*, 46.
49. Di Labio, G. A.; Hurley, M. M.; Christiansen, P. A. *J. Phys. Chem.* **2002**, *116*, 9578.
50. Geerke, D. P.; Thiel, S.; Thiel, W.; van Guststeren, W. F. *J. Chem. Theor. Comput.* **2007**, *3*, 1499.
51. Illingoworth, C. J. R.; Gooding, S. R.; Winn, P. J.; Jones, G. A.; Ferenczy, G. G.; Reynolds, C. A. *J Phys Chem A* **2006**, *110*, 6487.
52. Zhang, Y.; Lin, H., Truhlar, D. G.. *J. Chem. Theory Comput.* **2007,** *3*, 1378.
53. Zhang, Y.; Lin, H. *J. Chem. Theory Comput.* **2008**, *4*, 414.

Chapter 2

Enzymes

The enzymes are incredible molecular machines that make possible with extraordinary specificity and efficiency the happening of highly chemo-, stereo- and regio-selective reactions in less "drastic" conditions. Since their discovery, several theories have been proposed in order to understand, explain and predict enzymatic activity and its peculiarities. Enzymes are used in several fields of research and at different levels of production: are used in biofuel and paper industry, food processing and them inhibition/activation could be involved in treatments of several diseases. This chapter is a brief summary about principal aspects of enzymes world. Moving from a description of first theories proposed to the classification of the enzymatic architectures, will be introduced the research fields where they are involved, like catalysis and inhibition.

2.1 Introduction

Enzymes are macromolecules, mostly of protein nature, that function as catalysts increasing the reaction rate.[1] In fact are also known since 1980 the ribozymes, or RNA based enzymes. [2] Activities related to these supramolecular systems are known from the 17th century, but it occurred in the first years of 19th that scientists started to talk about "enzymes". [3,4] According with the reactions catalyzed by them, it is possible to distinguish six main classes, defining EC (Enzyme Commission): [5]

- EC 1, *Oxidoreductases* catalyze oxidation/reduction reactions;
- EC 2, *Transferases* transfer functional groups;
- EC 3, *Hydrolases* catalyze the hydrolysis reaction;
- EC 4, *Lyases* catalyze breaking bond reactions, different to hydrolysis and oxidation/reduction;
- EC 5, *Isomerases* favorite isomerization reactions;
- EC 6, *Ligases* catalyze bond formation reactions.

Moreover, in order to recognize the class of the enzyme, other numbers are added to *EC*, considering also the substrate nature and the reaction to which it is subjected.[5] Generally, an enzyme is characterized by a complex ternary and sometimes quaternary structure, but the protein surrounding has the most important aim to allocate the active site, the catalytic center where the reaction occurs. In the catalytic pocket, the substrate is adequately oriented, thanks to non-bonded interactions with amino acids mainly, to react with catalytic agents. Enzymes can accelerate reactions according with three verified ways:

1. Stabilizing the Transition States (TS); [6]
2. Providing alternative mechanism pathways; [7]
3. Destabilizing substrate ground state. [8]

Usually, the reactions in enzymatic environment happen as sum of all these contributions, with specific mechanistic strategies that will be discussed in more detail in next paragraphs.

2.2 One hundred years of models

The basic mechanism by which enzymes catalyze chemical reactions starts with the binding of the substrate (or substrates) to the active site on the enzyme. The active site is the specific region of the enzyme which binds the substrate. The binding of the substrate to the enzyme causes changes in the distribution of electrons in the chemical bonds of the substrate and ultimately causes the reactions that lead to the formation of products. The products are released from the enzyme surface to restore another reaction cycle. When scientists glimpsed about powerful prospects to study enzymatic reactions, several models were proposed. Representations of two enzymology milestones are depicted in **Figure 1**. The first, published in 1894, was the Fischer's "key-lock" model.[9] According whit this proposal, enzymes are rigid molecules (*lock*) that bind specifically in the active site (key hole) the correctly sized substrate (key), exactly as a key opens specifically a lock. (**Figure 1**)

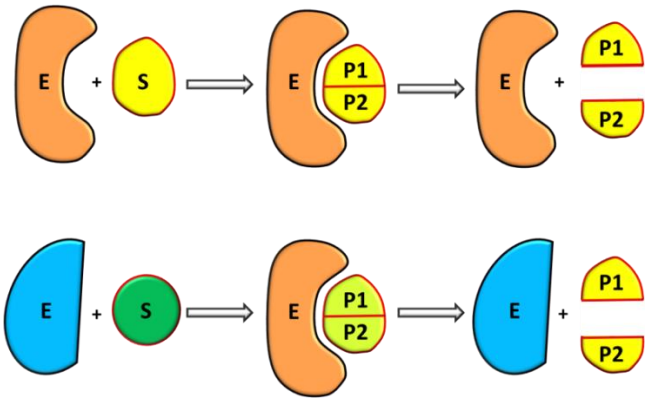


Figure 1. Schematic representations of "key-lock" model, proposed by Fischer (top), and "induced fit" model, proposed by Pauling (bottom).

With subsequent discoveries about secondary structures of enzymes and the increasing technology of experimental procedures to investigate them, [10] first Pauling [11] (1940) then Koshland [12] (1958) allowed to the description of enzyme-substrate ensemble with an "*induced fit*" model. [12] This is a modification of the *key-lock* model, since assumes that the substrate plays a role in determining the final shape of the enzyme and that the enzyme is partially flexible, modeling its active site around the substrate. This effect is driven from the mutual dynamic interactions between amino acid residues of the active site and the substrate. (**Figure 1**) This explains why certain compounds can bind to the enzyme but do not react because the enzyme has been distorted too much. Other molecules may be too small to induce the proper alignment and therefore cannot react. Only the proper substrate is capable of inducing the proper alignment of the active site. Despite to the great new different point of view, this theory looked to the reaction behavior considering mainly the

transition state stabilization effect. Nowadays, it is well known that enzyme-substrate interaction is a sum of complex events and stabilization of transition state may be part of them, as previously reported in **Scheme 1-Chapter 1.**

$$E + S \rightarrow ES \rightarrow EP \rightarrow E + P \qquad \text{(Eq. 2.1)}$$

2.2.1 Michaelis Menten Kinetic

The general theory of enzyme kinetics is based on the work of L. Michaelis and M. L. Menten,[13a] later extended by G. E. Briggs and J. B. S. Haldane. [13b] The basic reactions (E= enzyme, S= substrate, P= product) are shown in Eq. 2.2:

$$E + S \rightarrow ES \rightarrow E + P \qquad \text{(Eq. 2.2)}$$

Michaelis and Menten [13] obtained simple equations describing pseudo first order kinetics of enzyme-catalyzed reactions adopting simpler but similar approach to the **Scheme 1**. Assuming that the reactions are reversible and that a one-substrate enzyme catalyzed reaction is being studied, one can derive the Michaelis–Menten rate:

$$v = \frac{d[P]}{dt} = \frac{V_{\max}[S]}{K_M + [S]} = \frac{k_2[E]_0[S]}{K_M + [S]} \qquad \text{(Eq. 2.3)}$$

where V is the initial rate for first-order breakdown of the enzyme–substrate ([ES]) complex into enzyme (E) and product (P); V_{max} is the maximum reaction rate for a given concentration of enzyme in the presence of saturating levels of substrate; [S] is substrate concentration; and K_M is the Michaelis constant, the concentration of substrate required to achieve one-half the enzyme's maximal velocity. The equation applies to single-substrate reactions at a constant enzyme concentration. K_M is defined as:

$$K_M = \frac{k_{-1} + k_2}{k_1} \qquad \text{(Eq. 2.4)}$$

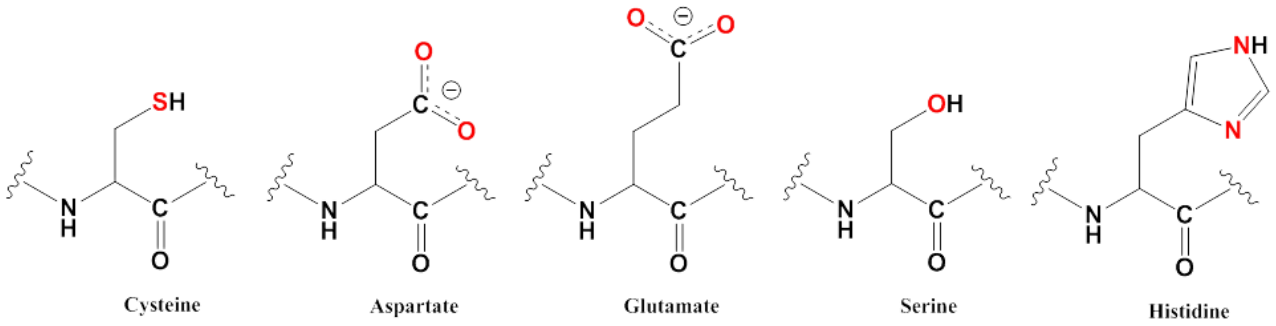
and represents a measure of the enzyme affinity to specific substrate. Going deeper, in the condition $[S] >> K_M$, setting k_2=k_{cat}:

$$v = V_{\max} = k_2[E]_0 = k_{cat}[E]_0 \qquad \text{(Eq. 2.5)}$$

The reaction rate is equal to the maximum rate, where k_{cat} is the overall rate constant.

2.3 Cofactors

Some enzymes require nonprotein groups, called *cofactors*, for their activity. An enzyme without its cofactor, called an *apoenzyme*, usually will be inactive in its catalytic role. An active enzyme with its cofactor is called a *holoenzyme*. Enzymes requiring metal ions as cofactors are called metalloenzymes. According with their nature, specific -R group of side chains of amino acid residues are the most common ligands observed in metalloproteins: example are imidazole of histidine, carboxylate of glutamate/aspartate, thiolate groups of cysteine and hydroxyl portion of serine. (**Scheme 1**)



Scheme 1. Representation of –R groups of the amino acids observed as metal ligand in metalloenzymes. In red, are highlighted the atoms where usually the bond with metal occurs.

Metal ions act with different contributes, depending specifically by the enzyme (and reaction catalyzed by them) and where they are allocated. They can be directly involved in catalytic mechanism [14] or show relevant contributions to the structural stability of protein [15] (being localized far from the active site); sometimes, without participating actively to the reaction, they can correctly orient the substrate facilitating the interaction with another catalytic agent, as a coenzyme or a water molecule. [16] It is estimated that 50% of proteins contain a metal ion: [17] well-known are systems containing Ca^{2+} and Mg^{2+},[18] ions of first transition metal period, like $Cu^{+/2+}$, Zn^{2+}, $Fe^{2+/3+}$, $Co^{2+/3+}$ and Ni^{2+}, [19] heavier ions like W species and even belonging to lanthanide period. [20] Coenzymes are other relevant catalytic agents, essential to the enzymes functionality. These organic cofactors are small molecules (comparing their dimensions to protein macromolecule), allocated in catalytic pocket, interacting tightly with amino acids surrounding and/or inorganic cofactors or covalently bonded to the enzyme.[21] The coenzymes, reacting with substrates, change their chemical nature, subsequently restored by other specific biological pathways, and this is the reason why they are considered a second substrate of enzyme.[22a] Examples are NAD^+/NADH, ATP, Coenzyme A, coenzymes found in all living cells, Thiamine Diphosphate (ThDP) and many others.[22] As for the inorganic cofactor, absence of coenzymes gives rise to serious diseases; moreover, these cofactors are not synthesized by human body and must be acquired from the diet.[23] In this thesis, the enzymes

undertaken to the mechanistic theoretical investigations, are examples of both metalloenzymes and organic cofactor containing ones.

2.4 Catalysis

Catalysis, in Chemistry, is one of the most famous concept. Generally, with this term are described all the processes involving in a reaction in presence of a *catalyst*, a chemical species able to accelerate the product formation. A good catalyst has a specific "fingerprint": it must be restored at the end of catalytic cycle, in order to react repeatedly, and must do no effect on the reaction thermodynamic, as evinced in **Figure 2**. Catalysts permit the reaction continuation with alternative pathways passing through intermediates and transition states with different energies, remarkably lower than the non-assisted reaction. Indeed one of the most important action of a catalyst is the energy stabilization of transition state, being the energy of that stationary point directly related to the rate of the reaction. [24] Higher is the activation energy, slower will be the reaction. Catalysis classification is based on the phases involved in the interaction catalyst-reactants: [25]

- Heterogeneous catalysis is defined when the catalyst and reactants phases are different;
- Homogeneous catalysis is observed when the species have the same physical phase;
- Enzymatic catalysis, sometimes considered as a special case between homogenous and heterogeneous catalysis, is defined, finally, when the catalysts are enzymes. Their celebrity belongs to the ability to increase the reaction rate, as previously mentioned. As biological machines, they make reaction happen with velocities requested by physiological time.

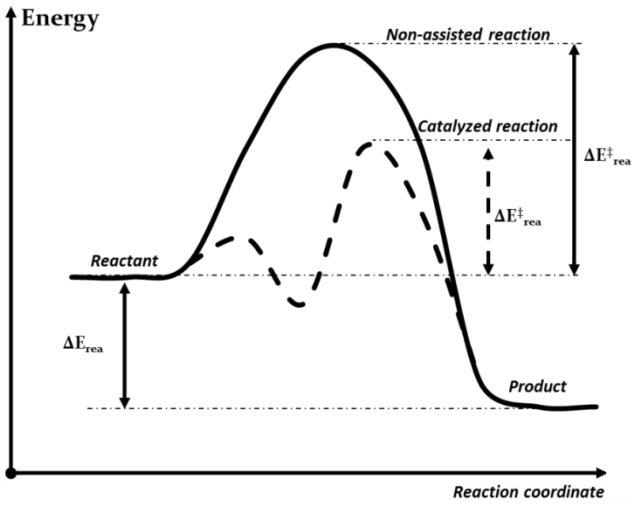
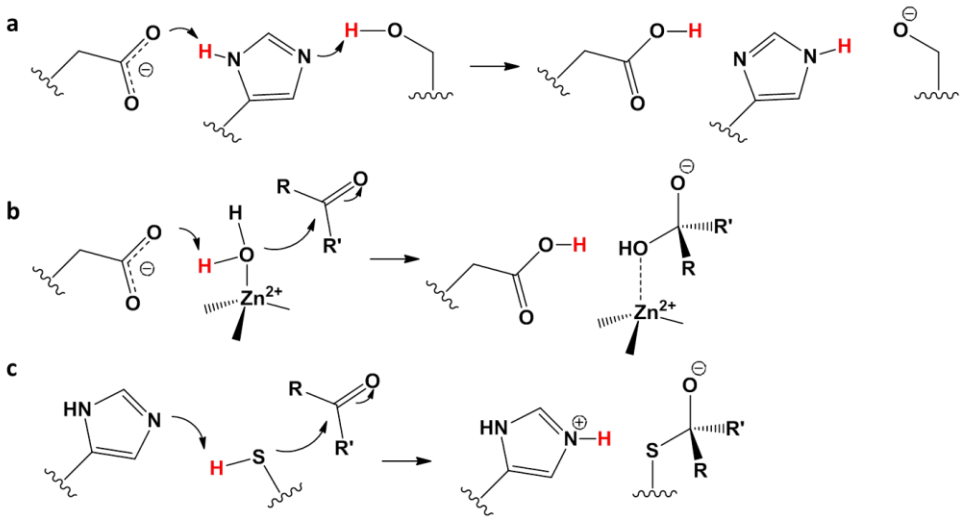


Figure 2. General schematization of a catalyzed (dot line) or non-catalyzed (black line) reaction.

Considering what was previously highlighted, the reaction mechanism in the enzyme active site happens thanks to specific interactions with the substrate. Usually, transition state stabilization, substrate ground state destabilization and the provider of alternative mechanism pathways are direct

consequences of three catalytic strategies: covalent, acid/base and electrostatic catalysis. [26] As far as the covalent catalysis is concerned, some enzymes can utilize nucleophilic amino acid sidechains (serine –OH, cysteine –SH and/or their ionized forms), or specific groups of coenzymes, to form covalent bonds with substrate. [27] General acid or base catalysis is observed when there are reactions occurring proton transfer. In this case, the presence of specific donor and acceptors agents, as the carboxylate group of aspartate or glutamate residues, or imidazole ring of histidine, enhances the mobility of hydrogens, facilitating then reaction. (see **Scheme 2**) Often, covalent and acid/base catalysis may occur in the same pathway. [28] Finally, the presence of charged sidechain groups, as positively charged lysine and arginine or negatively charged aspartate and glutamate, helps the catalysis, establishing dipole-dipole or ion-dipole interactions between substrate and active site population. [29] Despite the hydrophobic cavity, with low dielectric constant, water molecules may play a key role acting as catalytic agent. [29,30] It is important to underline that the catalytic effect must not be related only to one of the cited tactics. Certainly, exist examples where there are contributions more effective than others, but is likewise probable that the catalytic process is the sum of several strategies, originating several effects.

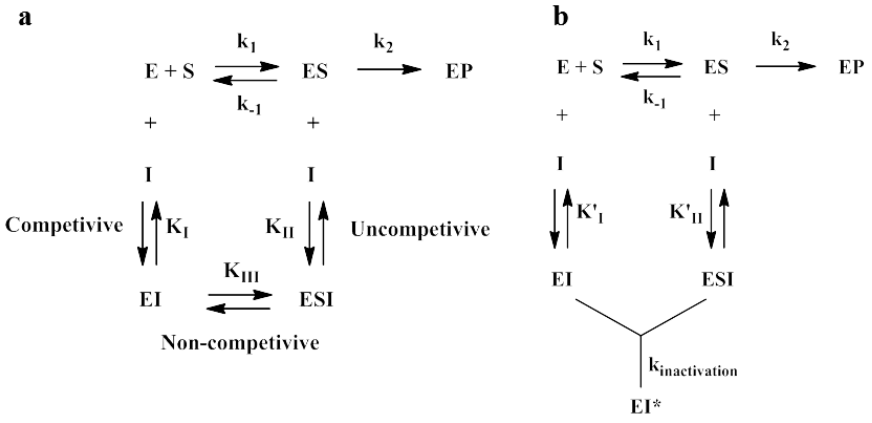


Scheme 2. Schematic representations of catalytic strategies in enzymatic catalysis: the acid/base (a), the electrostatic (b) and covalent (c) catalysis.

2.5 Inhibition

Enzyme inhibition, or more generically catalyst inhibition, is observed when the catalytic agent has not the possibility to process substrate molecules. This happens when an inhibitor, a molecule able to interact with the enzyme, induces different rearrangements of the enzyme structure. Different ways of inhibition are known, depending on the moment when enzyme-inhibitor complex (EI**)** forms and where the interaction takes place. [32] Generic scheme of the reversible inhibition is reported in **Scheme 3a**. In reversible inhibition, the inhibitor binds enzyme (competitive

mechanism) [32a] or the enzyme- substrate complex (uncompetitive mechanism) [32b] forming the enzyme-inhibitor (EI) or enzyme-substrate-inhibitor (ESI) complex, respectively.



Scheme 3. Representation of reversible (a) and irreversible (b) inhibition.

Noncompetitive inhibition is observed when a substrate molecule binds reversibly EI species. [32b] Finally, exist also the *mixed inhibition* that could be considered as a sum of all mechanisms previously showed. [32c] General irreversible inhibition action is shown in **Scheme 3b**. An irreversible inhibitor first forms a reversible complex EI or ESI, with the subsequent obtainment of EI* activated irreversible species. [33] Usually, these type of inhibitors, reversible too, may act with covalently modifications of active site, poisoning the catalyst; that effect is known as allosteric inhibition. [34] The case of glutathiol-S-transferase (**Paper IV**) is an example of uncompetitive inhibition.

2.6 *De Novo* proteins

De Novo protein world is a relatively new and fascinating research field. It is based on a basic rudiment: exist 20^x possible amino acids sequences for an x-residues protein, [35] taking into account for example, the 20 α amino acids. This means that knowing the fundamentals of protein folding should make possible to design a vast world of customized macromolecules, with activity specifically correlated. [36] The basis of that approach is fifty-years old [37] and stimulated researchers to go deeper the understanding of thermodynamics principles leading the folding process.[37] In the 2016,[37d] was firstly proposed the synthesis of an artificial *de novo* protein having catalytic function and in particular the esterase one. In detail, the catalytic triad Cys-His-Glu [37d] was installed in the secondary structure architecture of a seven-alfa helix chains protein. (**Figure 3**) The design of *de novo* functional biomolecules in which biostructure and function are built from scratch represents a considerable challenge. Enzyme-like catalysts obtained in this completely engineered way would have wide-ranging uses for manipulating and synthesizing small molecules with potential applications in medicine, industrial biotechnology and basic science. [38] With the aim to observe the

catalytic efficiency of the de novo protein (CC-Hept-Cys_{18}-His_{22}-Glu_{25}) a mechanistic study has been undertaken and the results have been compared with the experimental counterpart. (**Wp 3**)

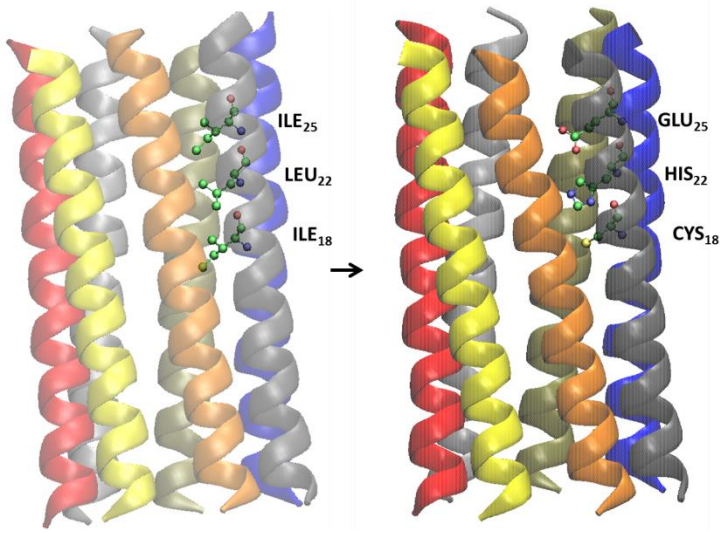


Figure 3. On the left, the starting architecture of protein adopted to install the enzymatic activity, substituting the triad Ile_{18}-Leu_{22}-Ile_{25} with the other one Cys_{18}-His_{22}-Glu_{25}.

References

1. IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997).
2. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell.* **1982**, *31*, 147.
3. Payen, A.; Persoz, J. F. *Annales de chimie et de physique* **1833**, *53*, 73.
4. Künhe, W. *Verhandlungen des naturhistorisch-medicinischen Vereins zu Heidelberg.* **1877**, *1*, 190.
5. Classification and Nomenclature of Enzymes by the Reactions they Catalyse, Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB).
6. Warshel, A; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. *Chem. Rev.* **2006**, *106*, 3210.
7. Cox, M. M.; Nelson, D. L. (2013). "Chapter 6.2. How enzymes work". Lehninger Principles of Biochemistry (6th ed.). New York, W.H. Freeman.
8. Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2003**, *301*, 1196.
9. Fischer, E. *Berichte der Deutschen chemischen Gesellschaft zu Berlin* **1894**, *27*, 2985.
10. Blow, D. *Structure* **2000**, *8*, R77.
11. Pauling, L. *Chem. Eng. News* **1940**, *24*, 1375.
12. Koshland, D. E. *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98.

13. a) Michaelis, L.; Menten, M. L. *Biochemische Zeitschrift* **1913**, *49*, 333. b) Lehninger, A. L.; Nelson, D. L.; Cox, M. M. (2000) Principles of Biochemistry, 3rd ed.,Worth Publishers, New York.
14. Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. G. *Nature* **2009**, *460*, 823.
15. Mikšovská, J.; Larsen, R. W. *Methods Enzymol.* **2003**, 360, 302.
16. Zheng, Y.-J.; Xia, Z.-X.; Chen, Z.-W.; Mathews, F. S.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 432.
17. Thomson, A. J.; Gray, H. B. *Current Opinion in Chemical Biology* **1998**, *2*, 155.
18. a) Brini, M.; Ottolini, D.; Calì, T.; Carafoli, E. (2013). "Chapter 4. Calcium in Health and Disease". Interrelations between Essential Metal Ions and Human Diseases. Metal Ions in Life Sciences. (A. Sigel, H. Sigel and R. K. O. Sigel. Eds.). Springer. b) Ryan, M. F. *Ann. Clin. Biochem.* **1991**, *28*, 19.
19. Barber-Zucker, S; Shaanan, B.; Zarivach, R. *Scientific Reports* **2017**, *7*, 16381 and references herein.
20. a) Kroneck, P. M. *J. Biol. Inorg. Chem.* **2016**, *21*, 29. b) Pol, A.; Barends, T. R. M.; Dietl, A.; Khadem, A. F.; Eygensteyn, J.; Jetten, M. S. M.; Op den Camp, H. J. M. *Enviromental Microbiol.* **2014**, *16*, 255.
21. Chapman-Smith, A.; Cronan, J. E. *Trends Biochem. Sci.* **1999**, *24*, 359.
22. a) Pollak, N.; Dölle, C.; Ziegler, M. *Biochem. J.* **2007**, *402*, 205. b) Hashim, O. H.; Adnan, N. A. *Biochemical Education* **1994**, *22*, 93.
23. a) Prasad, A. S. *British Med. J.* **2003**, *326*, 409. b) Guyatt, G. H; Patterson, C.; Ali, M.; Singer, J.; Levine, M.; Turpie, I.; Meyer, R. *Am. J. Med.* **1990**, *88*, 205.
24. Arrhenius, S. A. *Z. Phys. Chem.* **1889**, *4*, 96.
25. Rothenberg, G. (2008) Catalysis: concepts and Green Applications. Wiley-VCH Verlag GmbH & Co. KGaA
26. Kraut, D. A.; Carroll, K. S.; Herschlag, D. *Annu. Rev. Biochem.* **2003**, *72*, 517.
27. Storer, A. C.; Menard, R. *Meth. Enymol.* **1994**, *244*, 486.
28. Wilcox, P. E. *Meth. Enzymol.* **1970**, *19*, 64.
29. Fadouloglou, V. E.; Balomenou, S.; Aivaliotis, M.; Kotsifaki, D.; Arnaouteli, S.; Tomatsidou, A.; Efstathiou, G.; Kountourakis, N.; Miliara, S.; Griniezaki, M.; Tsalafouta, A.; Pergantis, S. A.; Boneca, I. G.; Glykos, N. M.; Bouriotis, V.; Kokkinidis, M. *J. Am. Chem. Soc.* **2016**, *139*, 5330.
30. Piazzetta, P.; Marino, T.; Russo, N. *Inorg. Chem.* **2014**, *53*, 3488.
31. Cleland, W. W. *Biochim. Biophys. Acta.* **1963**, *67*, 173.
32. a) Srinivasan, B.; Tonddast-Navaei, S.; Skolnick, J. *Eur. J. Med. Chem.* **2015**, *103*, 600. b) Cornish-Bowden, A. *Biochem. J.* **1974**, *137*, 143. c) Lee, L. G.; Whitesides, G. M. *J. Org. Chem.* **1986**, *51*, 25.
33. Maurer, T.; Fung, H. L. *AAPS pharmSci.* **2000**, *2*, E8.
34. Okar, D. A.; Lange, A. J. *BioFactors* **1999**, *10*, 1.
35. Huang, P-S; Boyken, S. E.; Baker, D. *Nature* **2016**, *537*, 320.
36. Epstein, C. J.; Goldberger, R. F.; Anfinsen, C. B. *Cold Spring Harb. Simp. Quant. Biol.* **1963**, *28*, 439.
37. a) Anfinsen C. B. *Science* **1973**, *181*, 223. b) Craig, E. A. *Science* **1993**, *260*, 1902. c) Caetano-Anollés, G.; Wang, M.; Caetano-Anollés, D.; Mittenthal, J. E. *Biochem. J.* **2009**,

417, 621. d) Burton, A. J; Thomson, A. R.; Dawson, W. M.; Brady, R. L.; Woolfson, D. N. *Nature chemistry* **2016**, *8*, 837.
38. a) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001**, *409*, 258. b) Nanda, V. *Nature chemistry* **2010**, *2*, 15. c) Hilvert, D. *Annu. Rev. Biochem.* **2013**, *82*, 447.

Chapter 3

Results

*This chapter is an overview about relevant aspects for every studied reaction in enzymatic environment. The first section is devoted to discuss about metalloenzymes, applying the pure QM methodology (**Paper I ,II** and **III**) and both pure QM and hybrid QMMM methods (**Paper IV**). The second and third Sections are dedicated to the study of an enzyme thiamine di-phosphate dependent (**Wp1**) and to the investigation of inhibition mechanisms of enzyme involving the glutathione (**Paper V**), adopting MD, pure and hybrid QM methodologies. Finally, mechanistic insights about a full de novo catalyst are presented. (**Wp2**)*

Section 1

3.1 Metalloenzymes

As mentioned in **Chapter 2**, metals in enzymes have different roles and if the heavy atom is localized in catalytic pocket promoting the catalysis the metalloproteins may be defined metalloenzymes. The accurate knowledge of their mechanisms represents a key step in the rational design of biomimetic systems[1] and new efficient drugs. Indeed, it was observed that metalloenzymes play a crucial role in diffusion of many human diseases and almost the 7% of approved drugs from 2013-2017 are metalloenzymes inhibitors.[2]

3.1.1 Cerium Methanol de-hydrogenase *(Paper I)*

Methanol Dehydrogenases (MDHs) are enzymes of the oxide-reductase class, which catalyze the oxidation of methanol and of other primary alcohols to the corresponding aldehydes (EC 1.1.2.7).[3-7] Usually, enzymes belonging to that family are metal-dependent; in particular, the Ca^{2+}- and Mg^{2+}-MDHs are well known and their catalytic mechanisms were widely studied. [8-15] Surprisingly, a recent work reported the discovery of *Methylacidiphilum fumariolicum* bacterium (XoxF-type) operating in extreme living conditions and its rare earth metals dependent MDH. (PDB:4MAE).[16] The x-ray structure evidenced a heavy atom directly bonded to the pyrrolo-quinoline quinone (PQQ) cofactor in the active site. Due to the presence of high Ce concentration, the electron density was supposed and assigned to cerium cation. [16] This supposition was confirmed by growth study evidences: in absence of rare earth metals the bacterium did not show growth, despite the presence of common Ca^{2+} and Mg^{2+} cations. [16] The scientific community looks with great interest to the research about methanol oxidation, being the second step of methane oxidation cycle, an important multi-step process, considered next-generation carbon feedstock.[17] Furthermore, the discovery of biological relevance of REMs opens to new scenario, regarding the adopting of these heavy atoms

species in different research fields. The aim of the work was to evaluate the effect of rare-earth metal presence to the catalytic mechanism, in comparison with Ca-MDH, previously studied,[10] applying the cluster approach.

3.1.1.1 Model and catalytic mechanism

The catalytic site is characteristic for the presence of Ce^{3+} with coordination number C.N.=9, as usual for lanthanides species.[18,19] The coordination shell presents Glu172, Asn256, Asp299 and Asp301 directly bonded to central metal. The PQQ cofactor completes inner coordination sphere. The cofactor is also involved in several H-bonds with other amino acids: Glu55, Arg110, Ser169, Arg326 and Asp388. All that residues were considered in the final cluster model (113 atoms including methanol substrate) used for the simulations, as depicted in **Figure 1a**. The investigated mechanisms are shown in **Figure 1b**. The first part (ES→TS1→INT1) represents the common step of the reaction. After this, the reaction follows two different pathways, proceeding through elimination-protonation (A) or retro-ene (B) reactions. The mechanism B showed energetic profile incoherent with the enzymatic catalysis requirements and will not be discussed in this thesis.

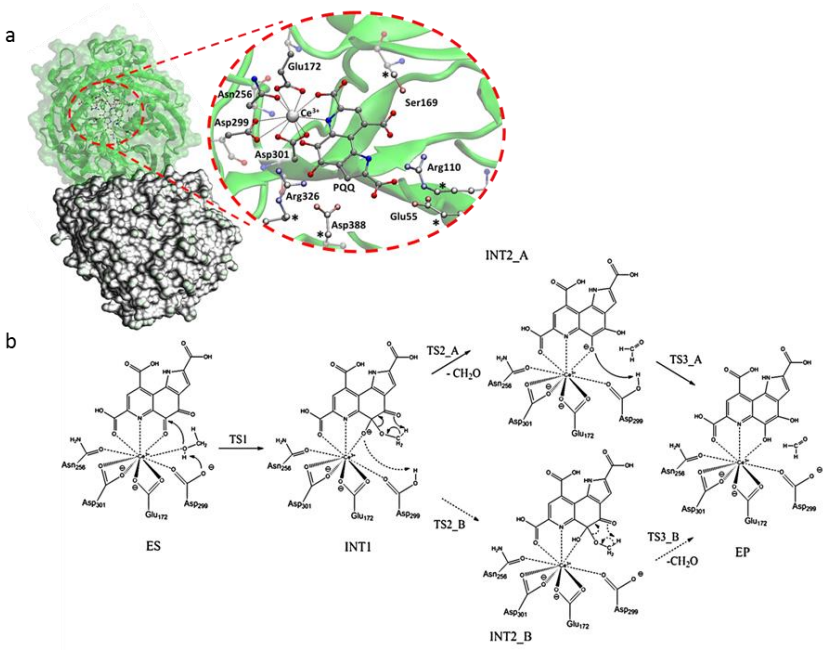


Figure 1. (a) In the red circle, the active site model adopted to study Ce-MDH reaction and its schematic representation. Atoms labelled with "*" were kept frozen during the optimization. First coordination and second shells are depicted with ball and sticks representations. (b) Proposed Ce-MDH reaction mechanisms (Scheme adapted from Paper I).

3.1.1.3 Comparison of catalytic behavior of Ce-MDH with that of Ca-MDH

The PESs of Ca-MDH and Ce-MDH are shown in **Figure 2a**. The comparison remarks that the effects of the different chemistry of the two metal ions (Ca^{2+} and Ce^{3+}) cover the first step of the

mechanism. In particular, the TS1 barrier is 14.0 kcal mol^{-1} and 6.5 kcal $mol^{-1,}$ for Ca-MDH and Ce-MDH, respectively. A possible explanation to that difference was found analyzing the NBO analysis performed on ES complex. The histogram (**Figure 2b**) reveals a more negative charge localized on the substrate, in Ce-MDH, due to the greater Lewis acidity, with respect to Ca-MDH. In fact the charge distribution makes the OH substrate more negative charged and consequently the attack on PQQ (TS1) can occur more easily. The TS2 remains the highest energy barrier in both enzymes (23.0 kcal mol^{-1} and 19.0 kcal mol^{-1} for Ce-MDH and Ca-MDH, respectively), even though the presence of REM, suggesting that the main contribution to the catalysis, by the cation, occurs in first part of reaction while the second ones is led by the cofactor PQQ. More detailed information are included in the attached **Paper I**.

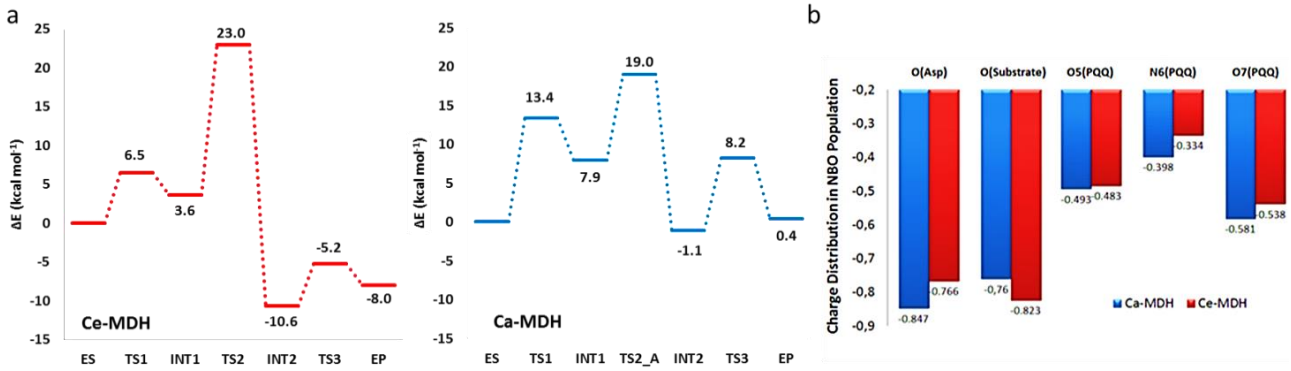


Figure 2. (a)The PESs calculated for Ce-MDH (red line) and Ca-MDH[10] (blue line) at ε=4/B3LYP-D3/6-311+G(2d,2p) level of theory and calculated charge distributions for ES complex (b). Figures have been adapted from Paper I.

3.1.2 Nitrile Hydratase containing Iron and Cobalt (Paper II)

The organic nitriles are widely adopted by industries in different production fields, like synthesis of plastic, pesticides, water treatment reagents and other pharmaceutical and chemical products. [30-31] In particular, nitriles are important starting reagents in the production of amides, adopted as drugs in tuberculosis [32] and other diseases. [33] On the other hand, industries product huge waste amounts containing these molecules, generally difficult to dispose, [34] that are toxic and have carcinogenic and teratogenic properties. [35-40] According with those requests, microbiological synthesis, mediated by biological catalysts, plays a key role in organic nitrile reactions, [41] and is adopted in industrial production since 1980. [42] Enzymes that catalyze the hydrolysis of organic nitriles are divided into two main classes: nitrilases and nitrile hydratases (NHases), [30] according to the nature of final product: nitrilase hydrolyzes the nitrile in corresponding carboxylic acid and ammonia; the nitrile hydratase converts nitrile in the amide. NHases, differently from nitrilases, present an inorganic cofactor in the active site, in particular Fe^{3+} and Co^{3+}.[43,44] One of the peculiarities is that Fe-type and Co-type NHases are rare biological example of iron and cobalt dependent system with non-heme and non corrinoid system, respectively,[45] that work at low spin multiplicities. The purpose of the work was to study the reactivity of Fe(III)-NHase and Co(III)-NHase with their natural

substrates,[46] aliphatic (pivalonitrile) and aromatic (benzonitrile) nitriles, respectively, including explicit water molecules, supposed and experimentally demonstrated important to the catalysis, [47] applying cluster approach.

3.1.2.1 Models

The starting models were built up by deposited x-ray crystallographic structures: Co(III)-NHase, from *Pseudonocardia Thermpophila* JCM 3095, with PDB code:1IRE [43] and Fe(III)-NHase, from *Rhodococcus erythropolis* N771, with PDB code: 2ZPE. [44] Both cations show a "claw setting" coordination sphere with an axial cysteine thiolate and labile water molecule, two equatorial peptide nitrogens and two equatorial sulphurs belonging to sulfenic (Cys-SOH) and sulfinic (Cys-SO_2H) cysteines. Despite the similarity of the two active sites, some differences are present in outer shell coordination sphere, attributable to the different substrate processed. Final cluster model selected for Co-NHase contains 116 atoms and a total charge of +1. The selected amino acids belong to both α and β subunits: αCys108, αCys111, αSer112, αCys113, bonded to Co^{3+}, and βArg52, βArg157, βLeu48, βTyr68, in the second shell. Two water molecules, present in the crystal structure (see **Figure 3a**) and the benzonitrile complete the atom-count of the model. In the Fe-NHase case, were included Tyr37, Tyr72, Tyr76, Arg56, Arg141 belonging to β subunit and Trp117 and Gln90, belonging to the α one, in addition to first shell sphere around Fe^{3+}, two water molecules and pivalonitrile substrate. The final model has 164 atoms and a total charge equal to +1.In both cases, the residue protonation state have been considered according to experimental observations. [17] The final models are reported in **Figure 3b**.

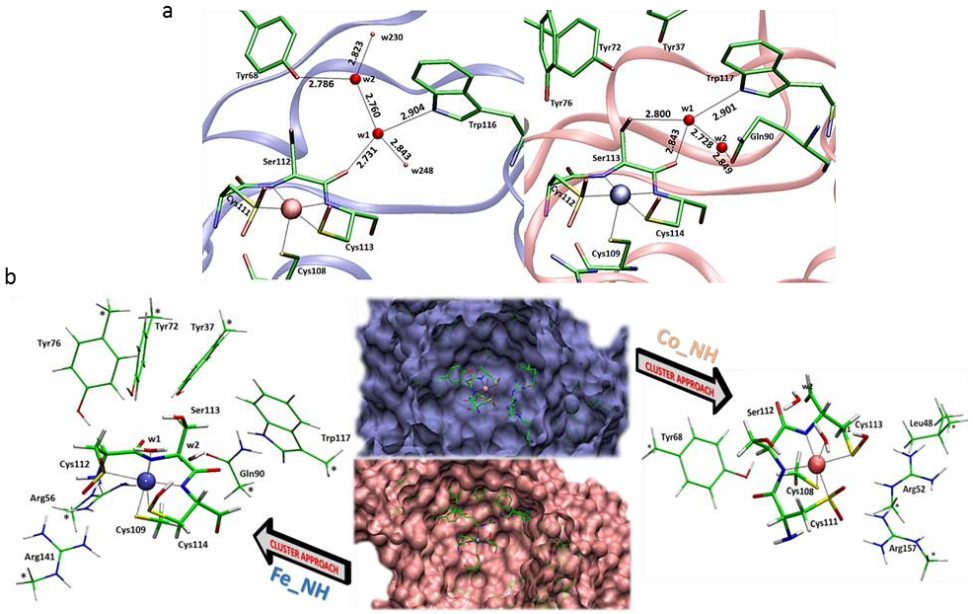


Figure 3. (a) X ray hydrogen bonds network for water molecules explicitly considered in the model for the Co-NHase and Fe-NHase, on the left and on the right respectively. (b) Active site models adopted to study metal-dependent NHases reaction and its schematic representation. Atoms labelled with "*" were kept frozen during the optimization. Figures have been taken from Paper II.

3.1.2.2 Reaction mechanism and comparison between Fe(III)- and Co(III)NHases

The proposed mechanism (**Figure 4a**) is based on several experimental and computational works on Fe-NHase. [16,18-21] On the other part, this study represents the first complete theoretical work about catalytic mechanism of Co-NHase, despite the presence of experimental observations. [14,18,19,22,23] As evidenced by PESs in **Figure 4b**, the rate determining is localized in last part of reaction (INT2-TS3) for both enzymes, with energy barriers estimated 26.3 kcal mol^{-1} and 24.6 kcal mol^{-1}, for Co- and Fe-NHase, respectively, accordingly with previous kinetics evidences. [16,18,19,24,25]

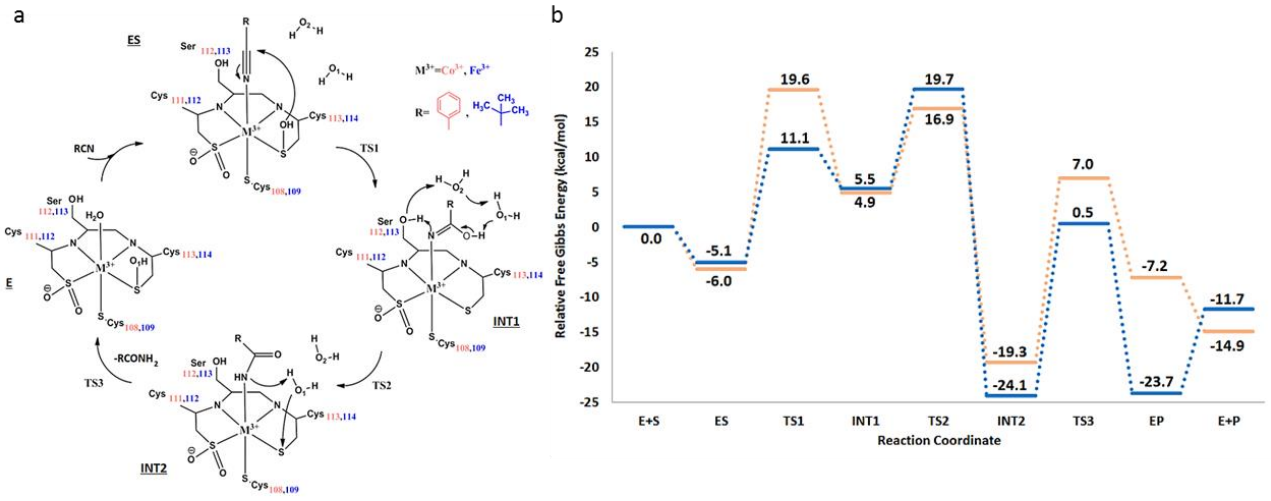


Figure 4. (a) Proposed mechanism for the nitrile hydrolysis catalyzed by Co^{3+}- and Fe^{3+}-NHase enzymes and (b) related PESs for reactions catalyzed by Co(III)-NHase enzyme (orange line) and by Fe(III)-NHase enzyme (blue line) obtained at M06L/6-311+G(2d, 2p) level and in protein environment (ε = 4). Scheme and figure have been adapted by Paper II.

In the Co-NHase case, the TS1 barrier (25.6 kcal mol^{-1}) is energetically comparable to the TS3 and sensibly higher than the same calculated for the Fe-NHase. A possible explanation to this evidence was found analyzing the TS1 (**Figure 5a**) and comparing electronic properties of ES complex (**Figure 5b**). Considering the optimized TS1 geometries, the C_{sp}-OH distances are 2.015 Å and 1.786 Å with bond orders of 0.239 and 0.305, in Co-NHase and Fe-NHase respectively. Furthermore, analyzing MO energy diagram is observed the of energy of the HOMO β and LUMO β in the ES complex of Fe(III)-NHase decreases much more than that of the same orbitals in the Co(III)-NHase. This means that a better overlap occurs between the HOMO β of free enzyme with the HOMO of nitrile with a consequent stronger covalent interaction. Accurate descriptions and explanations are reported in attached manuscript **Paper II.**

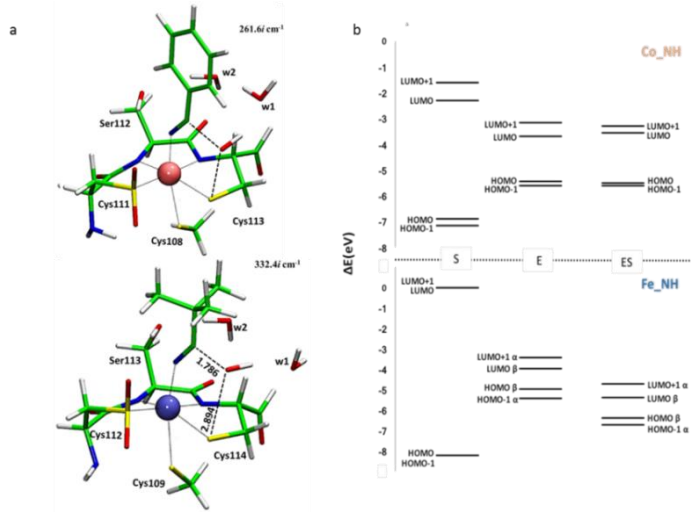


Figure 5. (a) TS1 optimized geometries of Co- (top) and Fe-NHases (bottom). (b) Frontier orbitals displacements of substrates, enzyme and enzyme complex systems (from left to right). The figures and the diagram have been taken from Paper II.

3.1.3 Peptidoglycan N-acetylglucosamine deacetylase containing Zinc (Paper III)

Nature exploits different strategies for enhancing the catalytic activity of enzymes resorting to beneficial mutations. The case of the proline hydroxylation in active site of polysaccharide deacetylase from Bacillus Cereus Bc1960 is an interesting example as small chemical modifications causes significant improvements in enzymatic activities. Starting from preliminary Molecular Dynamics simulations on the enzyme-substrate complex we have undertaken QM cluster investigation to explore the deacetylation mechanism considering both OH-proline and proline amino acids in the enzyme active site. Our calculations highlight as the hydrogen bond network established by the –OH group on the Cα of the proline with its closer amino acid neighbours stabilizes the transition states and, consequently, the occurring of the reaction. The analysis of the obtained PESs reveals as also the intermediate and the product get energetic benefits from the presence of hydroxyl-proline. These results corroborate the experimental findings on the polysaccharide deacetylase members that propose a higher deacetylase activity by a factor of 10 with respect to the non-hydroxylated form.

3.1.4 5-carboxyvanillate decarboxylase containing Manganese (Paper IV)

With the aim to give useful contributions in the computational enzymology panorama, about the preferred use of pure QM or hybrid QMMM approach, the present study has been undergone on a Mn^{2+}-dependent enzyme: the LigW Decarboxylase. This metalloprotein, that belongs to the superfamily of amidohydrolase (AHS), catalyzes the non-oxidative de-carboxylation of 5-carboxyvanillate (5-CV). The released products are vanillate and one molecule of CO_2, as usual in nature of several biological mechanisms.[14] The 5-CV is one of the last product of bi-phenyl lignin derivatives degradation, a relevant process from commercial and biotechnological point of view. Lignin degradation is a crucial step of plant biomass conversion in renewable aromatic chemical

and biofuels. [15] In the last two years, accurate computational and experimental studies were published on that enzyme, establishing the catalytic mechanism and the nature of released products.[16-17] These works have been devoted to clear mechanistic aspects of LigW catalysis. Primarily, was demonstrated that the de-carboxylation product is CO_2 and not $HCO3^-$, applying membrane inlet mass spectrometry (MIMS) and a 304 atoms QM cluster model, [17]computationally onerous but accurate. Intrigued by this last result, three different size models have been adopted, considering full QM cluster and hybrid QMMM methodologies, selecting medium size QM region. The Density Functional Theory (DFT) was selected as *ab initio* method for QM region, in each model, evaluating the efficiency of three different functionals: the pure M06L [18], the meta-hybrid MPWB1K [19] and the hybrid B3LYP. [20,21]

3.1.4.1 Models setup and reaction mechanisms

In the adopted three dimensional structure (PDB 4QRN), the active site pocket shows an hexa-coordinated Mn^{2+} ions directly bounded to the 5-nitrovanillate (5-NV) , a substrate-like inhibitor, which occupies two equatorial bound sites. The coordination sphere of metal cation is completed by one water molecule (w1) and Glu19, in equatorial position, and Asp314 and His158 in axial displacement The active site presents other residues directly interacting with the substrate: two Arginine, 58 and 252, involved in hydrogen bonds with carboxylate portions of molecule, Phe212 and Tyr317, above and below substrate and His241. Furthermore, two water molecules are presents in catalytic pocket, as shown in **Figure 6**. [16] The cluster model has 126 atoms, all treated at DFT level of theory. The ONIOM 1 model was obtained considering 15 Å of atoms surrounding the QM region, with a final size of 2154 atoms. An outer sphere with a thickness of 2 Å was maintained fixed during optimizations, while the remaining was let free. The ONIOM 2 model was the biggest created system including the whole enzyme and a sphere of water molecules, around the catalytic pocket. The final model consists of 11895 atoms. During optimizations, the water molecules and enzyme atoms far more than 18 Å from high layer region were maintained fixed in their positions. In **Figure 6b** are shown the studied reaction mechanisms, considering the formation of CO_2 (I)and HCO_3^-(II).

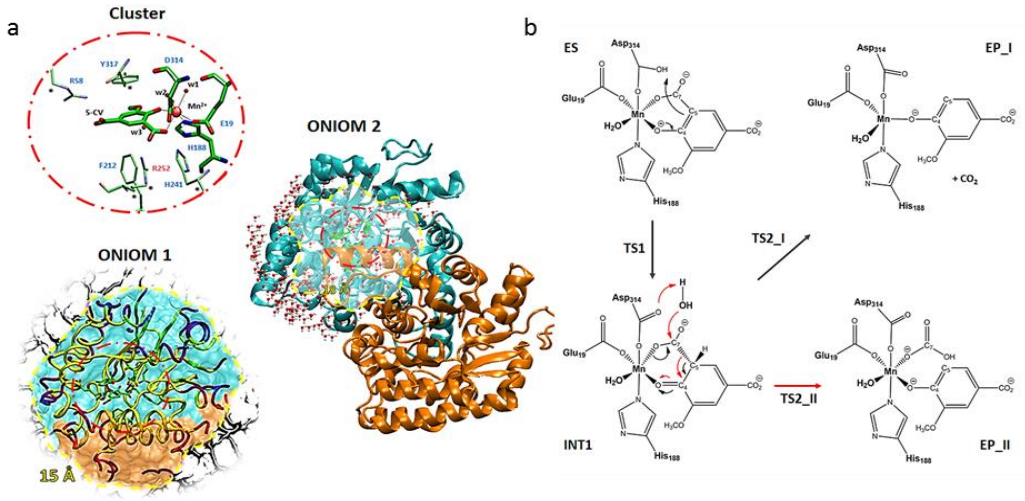


Figure 6. (a) Representations of three models adopted in the study and (b) proposed mechanisms. Figure and scheme have been taken from Paper IV.

3.1.4.2 Effect of models on PES

All the models, adopting the three different DFT functionals, proposed the TS1 barrier as the rate-limiting step. (**Table 1**) Based on the experimental k_{cat} value of 27 s^{-1} for *Sphingomonas paucimobilis* LigW, [17] the activation barrier requested by the system is close to the value of 16 kcal mol^{-1}. The closeness of TS1 values is appreciable, considering the functionals adopted in each model: average energy barriers of $\overline{\Delta E^{\ddagger}}$= 17.21 kcal mol^{-1}± 0.63 kcal mol^{-1}, $\overline{\Delta E^{\ddagger}}$= 14.42 kcal mol^{-1}± 1.03 kcal mol^{-1} and of $\overline{\Delta E^{\ddagger}}$= 17.04 kcal mol^{-1}± 1.36 kcal mol^{-1} have been calculated in the Cluster, ONIOM 1 and ONIOM 2, respectively.

	Cluster			ONIOM 1			ONIOM 2		
	SMD(ε=4)6-311+G(2d,2p)|SDD			6-311+G(2d,2p)|SDD: amber			6-311+G(2d,2p)|SDD:amber		
	B3LYP-D3	mPWB1K	m06L	B3LYP-D3	mPWB1K	m06L	B3LYP-D3	mPWB1K	m06L
ES	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TS1	16.33	17.70	17.61	14.29	13.23	15.74	14.69	12.15	15.28
INT1	10.55	9.62	7.85	5.26	7.19	3.98	7.39	7.29	7.89
TS2_I	15.13	15.07	14.58	15.08	11.89	14.49	13.45	12.63	14.00
EP_I	-3.42	-2.71	-4.48	5.40	8.99	4.71	-0.25	-6.14	-1.82
TS2_II	34.24	31.81	41.62	43.92	37.79	46.27	30.30	28.95	31.91
EP_II	3.86	8.89	-7.86	12.57	12.10	11.71	-6.86	-8.00	-4.78

Table 1. Energy values of all the species present on the potential energy surfaces of the two explored paths. The table has been adapted from Paper IV.

Closeness of geometrical parameters is observed even analyzing the optimized geometries of TS1. (**Figure 7**) The distance H_{Asp314}-C5 decreases to the values of 1.350 Å, 1.217 Å and 1.303Å, in the three models, respectively. The C5 carbon moved from sp^2 to the sp^3 hybridization and subsequently, the -COO^- portion of the substrate collapsed out of the phenyl ring plane. In the case of Cluster model this out of plane bond value is 19°, while in ONIOM 1 and ONIOM 2 is 47° and 76°, evidencing a more advanced situation detected by the hybrid models; these could be a possible explanation for the small mismatching between the energy barriers.

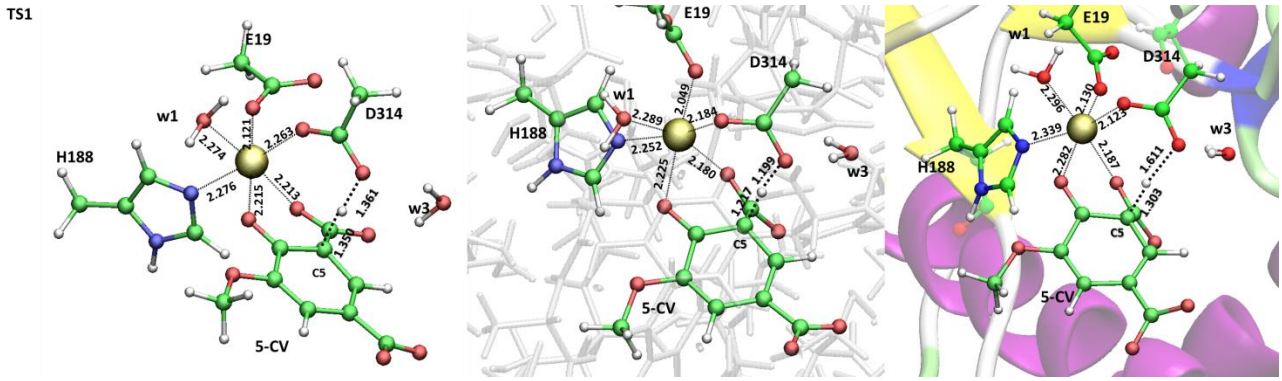


Figure 7. Optimized geometries of TS1 obtained for Cluster (left), ONIOM 1 (middle) and ONIOM 2 (right) models. The figures have been taken from Paper IV.

Section 2

3.2 human Transketolase (Wp 1)

The origin of the catalytic power of enzymes is a subject under debate for decades. It has been proposed that proficient enzymes not only stabilize the rate-determining transition states but also preclude intermediates to become too shallow as to keep the barriers of subsequent steps within an acceptable height. Transketolase (TK) was pointing out as a prototype of the latter. TK is a ThDP-dependent enzyme that catalyzes the transfer of a dihydroxyethyl group from the ketose D-xylulose-5-phosphate (X5P) to the aldose D-erythrose-4-phosphate (E4P), to yield the products D-fructose-6-phosphate (F6P) and D-gliceraldehyde-3-phosphate (G3P). Atomic resolution X-ray crystallography has shown that a crucial intermediate of the catalytic cycle displays a significant, putatively highly-energetic, out-of-plane distortion in a sp2 carbon of a lytic bond, postulated to lower the barrier of the subsequent step and to contribute for the increase the enzyme's turnover Here we use high-level QM/MM calculations to decipher the TK reaction mechanism, and to show that the origin of the distortion is not steric strain but mostly the establishment of an internal hydrogen bond in the intermediate, whose enthalpy pays-off for the distortion penalty necessarily to its alignment. The distortion associated to hydrogen bonding is present even in small gas-phase models, and in fact its net effect is stabilizing and anti-catalytic. The results help to understand the intrinsic enzymatic machinery behind enzyme's amazing proficiency.

Section 3

3.3 Inhibition of glutathione S-transferase (Paper V)

Piperlongumine (PL) is an anticancer compound whose activity is related to the inhibition of human glutathione transferase of pi class (GSTP1) overexpressed in cancerous tumors and implicated in the metabolism of electrophilic compounds. In the present work, the inhibition mechanism of hydrolyzed piperlongumine (hPL) has been investigated employing QM and QM/MM levels of theory. The potential energy surfaces (PESs) underline the contributions of Tyr residue close to G site in the catalytic pocket of the enzyme. The proposed mechanism occurs through a one-step process represented by the nucleophilic addition of the glutathione thiol to electrophilic species giving rise to the simultaneous C-S and H-C bonds formation. Both the used methods give barrier heights (19.8 and 21.5 kcal mol^{-1} at QM/MM and QM, respectively) close to that experimentally measured for the C-S bond formations. (23.8 kcal mol^{-1})

Section 4

3.4 *De novo* protein (Wp 2)

The reaction mechanism of an artificial esterase obtained by engineering the functional catalytic triad (Cys-His-Glu) into a fully de novo designed coiled-coil homo-heptameric peptide assembly

(CC-Hept) is proposed on the basis of a combined molecular dynamics (MD) and hybrid quantum mechanics/molecular mechanics (QM/MM) investigation. The preliminary MD simulations have been performed on both unbound and bound to the model substrate (p-nitrophenyl acetate) protein, thus supporting the stability of de novo protein architecture. Two reaction pathways have been deeply analyzed at QM/MM level evidencing the rate determining step in agreement with the observed kinetics evidences. The role played by the water molecules and the catalytic triad has been highlighted at atomistic level. Our results should be useful for future developments of more selective and efficient engineered enzymes.

References

1. Simmons, T. R.; Berggren, G.; Bacchi, M.; Fontecave, M.; Artero, V. *Coord. Chem. Rev.* **2014**, *270*, 127.
2. Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M. *Chem. Rev.* **2018** doi: 10.1021/acs.chemrev.8b00201
3. Anthony C. *Subcell. Biochem.* **2000**, *35*, 73.
4. Ghosh, M.; Anthony, C.; Harlos, K.; Goodwin, M. G.; Blake, C. *Structure* **1995**, *3*, 177.
5. Xia, Z.; Dai, W.; Zhang, Y.; White, S. A.; Boyd, G. D.; Mathews, F. S. *J. Mol. Biol.* **1996**, *259*, 480.
6. Xia, Z. X.; He, Y. N.; Dai, W. W.; White, S. A.; Boyd, G. D.; Mathews, F. S. *Biochemistry* **1999**, *38*, 1214.
7. Zheng, Y. J.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11881.
8. Zheng, Y. J.; Xia, Z.; Chen, Z.; Mathews, F. S.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 432.
9. N. B. Idupulapati, D. Mainardi, *J. Phys. Chem. A* **2010**, *114*, 1887.
10. M. Leopoldini, N. Russo, M. Toscano, *Chem. Eur. J.* **2007**, *13*, 2109.
11. P. Hothi, M. J. Sutcliffe, N. S. Scrutton, *Biochem. J.* **2005**, *388*, 123.
12. Zhang, X.; Reddy, S. Y.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 745.
13. Reddy, S. Y.; Bruice, T. *Protein Sci.* **2004**, *13*, 1965.
14. Idupulapati, N. B.; Minardi, D. S. *THEOCHEM* **2009**, *901*, 72.
15. Pol, A.; Barends, T. R. M.; Dietl, A.; Khadem, A. F.; Eygensteyn, J.; Jetten, M. S. M.; Op den Camp, H. J. M. *Enviromental Microbiol.* **2014**, *16*, 255.
16. Hwang, I. H.; Lee, S. H.; Choi, Y. S.; Park, S. J.; Na, J. G.; Chang, I. S.; Kim, C.; Kim, H. C.; Kim, Y. H.; Lee, J. W.; Lee, E. Y. *J. Microbiol. Biotechnol* **2014**, *24*, 1597.
17. Murugesu, M.; Schelter, E. J. *Inorg. Chem.* **2016**, *55*, 9951.
18. Djinovic-Carugo, K.; Carugo, O. *J. Inorg. Biochem*. **2015**, *143*, 69.
19. Mascharak, P. K. *Coord. Chem. Rev*.**2002**, *225*, 201.
20. Payne, S. W; Wu, S.; Fallon, R. D.; Tudor, G.; Stieglitz, B.; Turner, I. M. Jr.; Nelson, M. J. *Biochemistry* **1997**, *36***,** 5447**.**
21. Kamble, A. L.; Banoth, L.; Meena, V. S.; Singh, A.; Chisti, Y.; Banerjee, U. C. *3 Biotech*. **2012**, *3*, 319.
22. Lin, S. H.; Chong, Z. Z.; Maiese, K. *J. Med. Food.* **2001**, *4*, 27.

23. Wyatt, J. M; Linton, A. E. (1998) “The industrial potencial of microbial nitrile biochemistry.” Cyanide Compounds in Biology. (Evered, D.; Harnett, S. eds.) Ciba Foundation Symposium; Wiley-Interscience: Hoboken, NJ; *Vol. 140*.
24. Leonard, A.; Gerber, G. B.; Stecca, C.; Rueff, J.; Borba, H.; Farmer, P.B.; Sram, R.J.; Czeizel, A.E.; Kalina, I. *Mutat. Res.* **1999**, *436*, 263.
25. Wyatt, J. M.; Knowles, C.J. *Biodegradation* **1995**, *6*, 93.
26. Woutersen, R. A. *Scand. J. Work. Environ. Health* **1998**, *24*, 197.
27. Delzell, E.; Monson, R. R. *J. Occup. Med.* **1982**, *24*, 767.
28. Czeizel, A.E.; Hegedus, S.; Timar, L. *Mutat. Res*. **1999**, *427*, 105.
29. Kidd, H.; James, D.R. (1998) The Agrochemicals Handbook. (3rd ed.) Royal Society of Chemical Information Services, Cambridge, UK.
30. Novak, W; Ohtsuka, Y.; Hasegawa, J.; Nakatsuji, H. *Int. J. Quantum Chem.* **2002**, *90*, 1174.
31. Asano, Y.; Tani, Y.; Yamada, H. *Agric. Biol. Chem.* **1980**, *44*, 2251.
32. Miyanaga, A.; Fushinobu, S.; Ito, K.; Wakagi, T. *Biochem. Biophys. Res. Commun.* **2001**, *288*, 1169.
33. Hashimoto, K.; Suzuki, H.; Taniguchi, K.; Noguchi, T.; Yohda, M.; Odaka, M. *J. Biol. Chem.* **2008**, *283*, 36617.
34. Kovacs, J. A. *Chem Rev.* **2004**, *104*, 825.
35. Yamada, H.; Kobayashi, M. *Biochem*istry **1996**, *60*, 1391.
36. Martinez, S.; Wu, R.; Sanishvili, R.; Liu, D.; Holz, R. *J. Am. Chem. Soc.* **2014**, *136*, 1186.
37. Stein, N.; Gumataotao, N.; Hajnas, N.; Wu, R.; Lankathilaka, K. P. W.; Bornscheuer, U. T.; Liu, D.; Fiedler, A. T.; Holz, R. C.; Bennett, B. *Biochemistry* **2017**, *56*, 3068.
38. Dey, A.; Chow, M.; Taniguchi, K.; Lugo-Mas, P.; Davin, S.; Maeda, M.; Kovacs, J. A.; Odaka, M.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J. Am. Chem. Soc.* **2006**, *128*, *533.*
39. Light, K. M.; Yamanaka, Y.; Odaka, M.; Solomon, E. I. *Chem. Sci* **2015**, *6*, 6280.
40. Stein, N.; Gumataotao, N.; Hajnas, N.; Wu, R.; Lankathilaka, K. P. W.; Bornscheuer, U. T.; Liu, D.; Fiedler, A. T.; Holz, R. C.; Bennett, B. *Biochemistry* **2017**, *56*, 3068.
41. Yamanaka, Y.; Kato, Y.; Hashimoto, K.; Iida, K.; Nagasawa, K.; Nakayama, H.; Dohmae, N.; Noguchi, K.; Noguchi, T.; Yohda, M.; Odaka, M. *Angew. Chem. Int. Ed.* **2015**, *54*, 10763.
42. Hopmann, K. H.; Guo, J.-D.; Himo, F. *Inorg. Chem.* **2007,** *46*, 4850**.**
43. Hopmann, K. H.; Himo, F. *J. Inorg. Chem.* **2008,** *22*, 3452**.**
44. Hopmann, K. H. *Inorg. Chem.* **2014**, *53*, 2760.
45. Noguchi, T.; Nojiri, M.; Takei, K.; Odaka, M; Kamiya, N. *Biochemistry* **2003**, *42*, 11642.

Concluding Remarks

This thesis comprises quantum chemical studies of some metallo-enzymes of different enzymatic classes. Mainly density functional theory for quantum chemical cluster approach and hybrid quantum mechanics/molecular mechanics methods (in some cases coupled with molecular dynamics simulations) have been used to investigate the reaction mechanisms.

- For methanol-de-hydrogenase (Paper I) and nitrile hydratase (Paper II), other than a detailed description of the entire reaction mechanism, the role of the amino acid residues surrounding the active site as function of the metal ion was established.
- Important deeper insights, useful to rationalize the better catalytic performance of a polysaccharide deacetylase containing a modified proline (2-hydroxyl proline) in its active site with respect to that containing the natural proline (Paper III), were obtained by comparison of their potential energy profiles. The role of the hydrogen bonds network in the catalytic pocket during the reaction mechanism is deeply analyzed.
- The catalytic mechanism of the decarboxylation of 5-carboxyvanillate by LigW, a manganese dependent enzyme, has been elucidated by using QM cluster and hybrid QM/MM methodologies. Moreover, a detailed kinetic analysis is provided devoting particular attention to the main modeling factors. (Paper IV)
- The inhibition mechanism of piperlongumine, a new potential anticancer drug, by human glutathione transferase of pi class overexpressed in cancerous tumors and implicated in the metabolism of electrophilic compounds, has been investigated employing QM and QM/MM methods. The potential energy surfaces (PESs) underline the contributions of Tyr residue close to G site in the catalytic pocket of the enzyme. (Paper V)
- High-level QM and QM/MM calculations were used to elucidate the reaction mechanism of Transketolase,. Deeper analysis on the origin of the distortion in the reaction intermediate whose destabilization affects enzymatic catalysis was performed. (Work in progress 1)
- The potential esterase catalytic activity of a fully de novo designed coiled-coil homo-heptameric peptide assembly (CC-Hept) having the functional catalytic tryad (Cys-His-Glu) in its “catalytic pocket” was investigated at QM/MM level of theory (Work in progress 2). Above all two reaction pathways have been deeply analyzed evidencing the rate determining step in agreement with the observed kinetics evidences. The role played by the water molecules was also explored.

Paper I

"How Can Methanol Dehydrogenase from Methylacidiphilum fumariolicum Work with the Alien Ce^{III} Ion in the Active Center? A Theoretical Study"

Mario Prejanò, Tiziana Marino, Nino Russo
Chemistry, A European Journal **2017**, *23(36)***,** 8652-8657.

doi: 10.1002/chem.201700381

ChemPubSoc Europe

DOI: 10.1002/chem.201700381

CHEMISTRY A European Journal Full Paper

Computational Chemistry |*Hot Paper*|

How Can Methanol Dehydrogenase from *Methylacidiphilum fumariolicum* Work with the Alien Ce^{III} Ion in the Active Center? A Theoretical Study

Mario Prejanò, Tiziana Marino,* and Nino Russo[a]

Abstract: Lanthanides are an example of nonbiogenic metal species and have been widely used in crystallographic and spectroscopic studies to probe Mg^{2+}/Ca^{2+} binding sites in metalloproteins by replacing the native cofactor. Recently, a methanol dehydrogenase (MDH) enzyme containing cerium ion in the active site has been isolated from *Methylacidiphilum fumariolicum* bacterium. With the aim to highlight as metal ion substitution can be reflected in catalytic mechanism, a comparative DFT study between Ca- and Ce-MDH has been undertaken. The obtained potential energy surfaces (PES), for two considered reaction mechanisms (named A and B), indicate mechanism A (addition-elimination and protonation processes) as the favored for both the enzymes and show as the barrier for the rate-determining step of Ce-MDH requires 19.4 kcal mol^{-1}.

Introduction

Methanol dehydrogenases (MDHs) are enzymes (belonging to the class of oxidoreductases) that efficiently catalyzes the oxidation of methanol or other primary alcohols to the corresponding aldehydes.[1–4] These enzymes are, generally, calcium-dependent and their reaction mechanism has been investigated in great detail.[5–12]

Recently, in the waters of acidic hot spring of a volcano in south Italy, has been discovered the *Methylacidiphilum fumariolicum* bacterium (XoxF-type) that operates at extreme living conditions and the structure of its MDH revealed the presence of rare earth elements (La, Ce, Pr, and Nd) rather than calcium.[13] Lanthanides are not essential for human life, though they are ubiquitous at relatively low concentrations.[14,15]

Due to their modest abundance, the related biological activity received modest attention. Although the interpretations concerning the functions and activities of lanthanoid elements in the human organism are still debated, convergent opinion is that trivalent lanthanum ions function as calcium analogs in the biological systems substituting Ca in many proteins,[16] including enzymes[17] and cell membranes.[18,19]

The X-ray structure of the XoxF-type methanol dehydrogenase (MDH) enzyme of the aforementioned bacterium revealed a heavy atom in the active site bound directly to the pyrroloquinoline quinone (PQQ) ligand co-factor.[13] Due to the high concentration of cerium (approximately 1.0 μM), the anomalous electron density was assigned as a cerium cation, and was successively confirmed by growth studies evidencing that *M. fumariolicum* SolV actively stored cerium.[20] Furthermore, this bacterium did not show growth in the absence of rare-earth metal salts in spite of the presence of metal ions normally implicated in the physiological processes such as calcium ion.

To the best of our knowledge, information on the reaction mechanism for cerium-dependent MDH (Ce-MDH), differently from the well-studied calcium-dependent MDH (Ca-MDH),[5–12] are not present in literature. For this reason, we have undertaken a detailed study on reaction pathway for this rare-earth-metal-dependent MDH considering the previous studies on the proposed mechanisms for Ca-MDH.[5–12]

This assumption is supported the differences between the two enzymes concern only one residue in the coordination shell of the metals.

Encouraged from the findings of the recent DFT work devoted to determine the electronic structure of the catalytic active site in Ce-MDH metalloenzyme[20] and from the previous investigation on the catalytic mechanism of Ca-MDH from the *Methylophilus methylotrophus* W3A1,[8] we have employed the quantum chemical cluster methodology to build up the potential energy surfaces relative to two possible reaction mechanisms promoted by Ce^{III}-dependent MDH enzyme.

Computational Details

All the computations have been carried out with the Gaussian 09 software package.[21]

The hybrid Becke exchange and the Lee, Yang, Parr correlation (B3-LYP)[22,23] functions have been used in the geometries optimization, considering the dispersions contribution[24] as implemented in Gaussian package.

[a] *M. Prejanò, Prof. T. Marino, Prof. N. Russo*
Dipartimento di Chimica e Tecnologie Chimiche
Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)
E-mail: tiziana.marino65@unical.it

Supporting information and the ORCID identification number for the author of this article can be found under https://doi.org/10.1002/chem.201700381.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

The 6–31+G(d,p) basis set has been chosen for the C, N, O, H atoms, whereas the SDD[25] pseudo-potential coupled with its related basis set has been used for the Ce. The same computational protocol was successfully followed in the mechanistic studies of other metal-proteins.[26–29]

For all stationary points, intermediates, and transition states (TS), frequencies calculation was performed at the same level as geometries optimization to confirm their nature of minima (no negative vibrational frequency) and transition states (only one negative vibrational frequency). Due to the nature of the computations performed with some constrained atoms, there were more relatively small negative frequencies that do not obscure the main negative TS-frequency.[26] To model the protein environment effects, the polarizable continuum technique based on C-PCM continuum model[30,31] was used. The solvation energies were obtained from single point calculations performed on the optimized geometries with the more extended basis set 6–311+G(2d,2p)|SDD. The choice of the dielectric constant $\varepsilon=4$ is usually considered to be a good representation of protein surrounding.[32–38] For the potential energy surfaces (PESs), the final solvation energies reported include the ZPE corrections. The unrestricted Kohn–Sham (UKS) calculations have been performed and small spin contaminations has been found, being the $\langle S2\rangle$ value equal to 0.77. We have considered the spin multiplicity ($2S+1=2$) of the cluster previously determined as the ground state.[20]

NBO analysis[39] was performed on all the intercepted stationary points of the investigated PESs.

Cluster Model: Our initial structure was modelled from X-ray structure of the Xoxf-type natural cerium-dependent MDH (PDB: 4MAE).[13] The QM cluster includes all the amino acid residues present in the inner coordination shell of Ce^{3+} ion (Glu172, Asn256, Asp299, Asp301), the cofactor 2,7,9-tricarboxypyrroloquinoline quinone (PQQ) and other residues of the outer coordination shell involved in H-bonds with the cofactor (Glu55, Arg110, Ser169, Arg326, Asp388).

The considered Asp and Glu residues are modelled by a CH_3COO^- group, the Ser169 by CH_3OH, the Asn256 by CH_3CONH_2 and the arginine residues by $CH_3NH(NH_2)_2^+$ as it can be evinced by Figure 1. Following the standard procedure in the modelling enzymes with the cluster approach[8,26,40–44] to keep the optimized structures resembling the experimental one, the truncation atoms except those of PQQ and metal-coordinated amino acids, have been fixed at their crystallographic positions during the geometries optimization. The final model consists of 113 atoms, including the substrate (CH_3OH) and has a total charge of zero. Since the first step of the reaction has been ascertained to be the protonation of Asp299, the substrate must be protonated.[5–7,9]

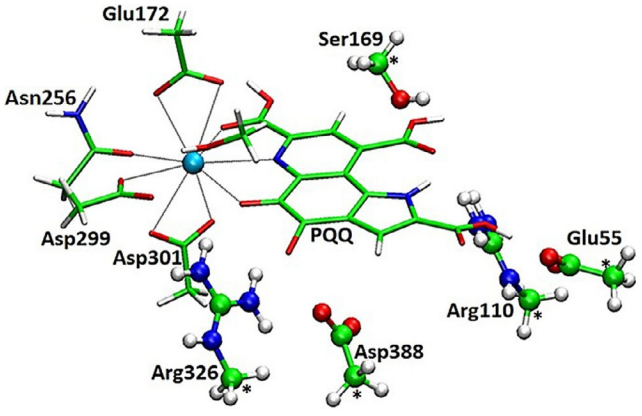


Figure 1. Model cluster representing the active site of Ce-MDH. Stars indicate the locked atoms kept frozen during the calculations.

Results and Discussion

The Ce-MDH active site has been studied employing DFT by Bogart and co-workers that have determined the cerium(IV/III)-PQQ^0 redox potential and assigned the oxidation state of cerium (+3) coupled with the doublet ground state spin multiplicity, as above mentioned.[20] On the basis of these findings, the authors propose the Ce-MDH as better catalyst than Ca-MDH.[20]

In the mechanism of methanol oxidation to formaldehyde by MDH, the cofactor PQQ is responsible of the 2 e^- reduction starting from its neutral quinone form to the dihydroquinone reduced one ($PQQH_2$).[5,20] A number of possible mechanisms have been proposed for MDH over the years[5,6,8–11,45–48] confirming as quino-proteins play an important role in the field of enzymatic catalysis.[4–10]

On the basis of a recent theoretical investigation on Ca-MDH,[8] we have considered as more probable the two pathways depicted in the Scheme 1.

Mechanism A involves the residue Asp299 as general base catalyzing the addition of methanol to the PQQ carbonyl group at C5 position with the formation of CH_2O and deprotonation of Asp299 in two distinct processes (addition–elimination and protonation).

Mechanism B suggests, starting from the first intermediate (INT1), a retro-ene reaction with the simultaneous formation of CH_2O.

Starting from these two proposed mechanisms and considering that the role of metal ion in MDH is not clear, we have undertaken a detailed theoretical analysis of the potential energy surfaces for cerium containing MDH by using density functional theory and by using a cluster model similar to that previously used in the calcium dependent MDH.[8]

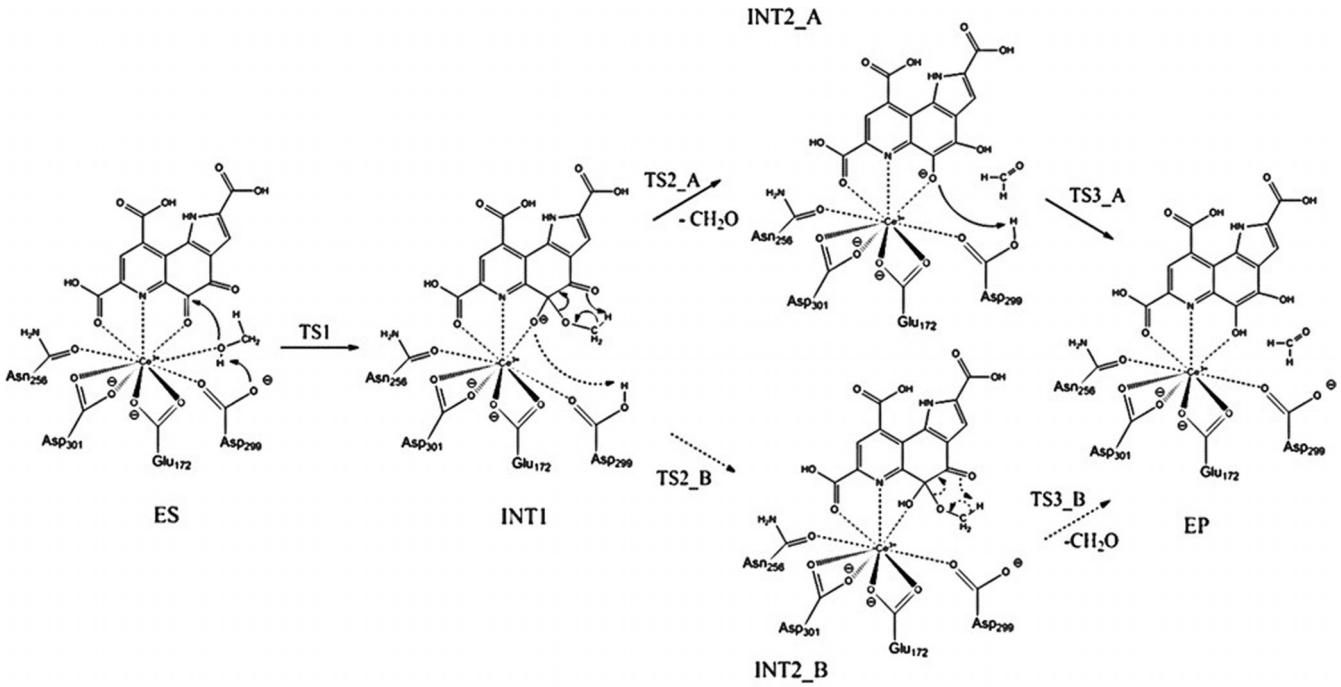
The reliability of cluster model in enzymatic catalysis has been largely explored in the last decades by different groups and for a huge number of enzymes.[26,33,40–44,49]

As previously mentioned the two considered mechanisms share the same behavior from the ES to INT1 species but going forward the mechanism B proceeds with higher energetics.

Due to the presence of a barrier incompatible with an enzymatic catalytic activity (61.8 kcal mol^{-1} for the rate-determining step described by TS3_B of Scheme 1) our data exclude the possibility that the reaction occurs throughout a retro-ene reaction with the simultaneous formation of formaldehyde product. For this reason, we will discuss in detail only the PES related to the mechanism A (shown in Figure 2), whereas that of the alternative one (B) is given in Figure S1 in Supporting Information.

All the optimized geometries of the stationary points present on PES relative to mechanism A are depicted in Figure 3.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 1. Two possible mechanisms of methanol oxidation catalyzed by the Ce-MDH considered in the present study.

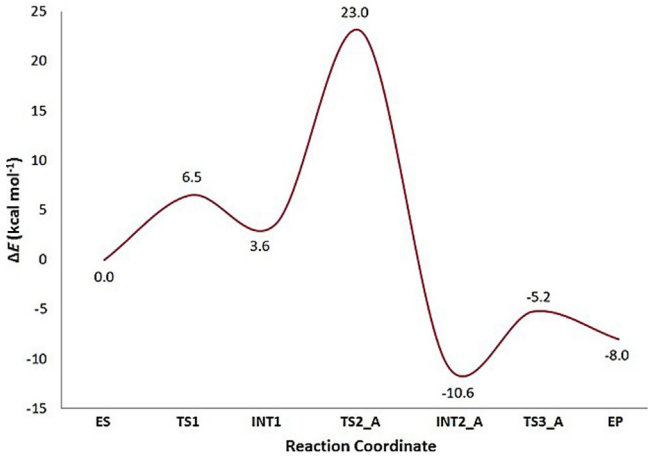


Figure 2. Calculated solvation energy B3LYP-D3/6–311 + G(2d,2p) | SDD for the addition—elimination–protonation (mechanism A) in protein environment ($\varepsilon = 4$).

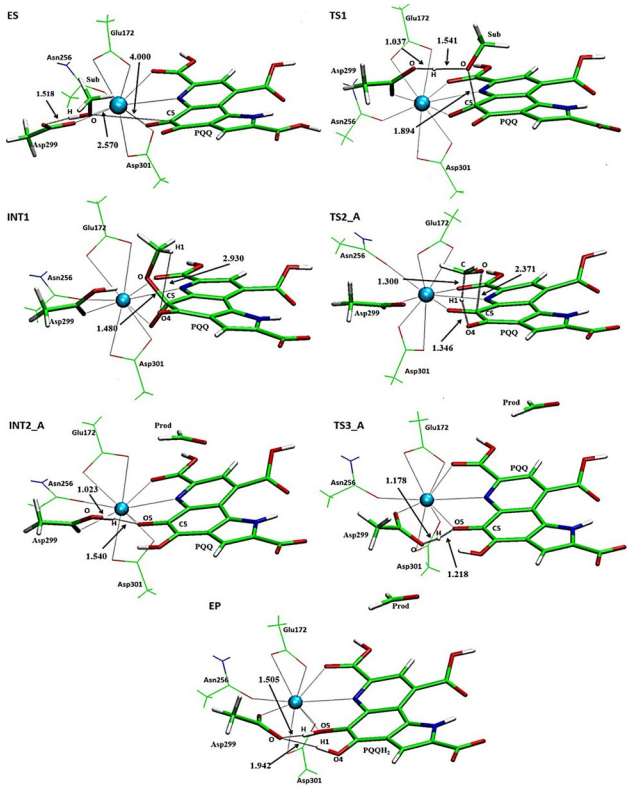


Figure 3. Optimized geometries of the stationary points (mechanism A) at B3LYP-D3/6-31G + (d,p) level of theory. Distances are given in Å. For clarity, the amino acid residues of the outer coordination shell are omitted. For the full model see Figure 1.

The optimization of the enzyme-substrate complex (ES) generates a structure whose geometrical parameters are in well agreement with the experimental counterpart.[13]

The cerium shows a coordination number (CN) equal to 9 as generally occurred in the lanthanides chemistry.[50,51]

The PQQ acts as bidentate ligand due to the involvement of O7 (2.730 Å) and O5 (2.960 Å) atoms. This coordination contributes to the polarization of the C5–O5 carbonyl group in which the C5 atom becomes more positive (0.53 e) with respect to that of the PQQ alone (0.46 e) due to the electron-withdrawing effect of the metal ion as indicated by NBO analysis, and prepares for the next nucleophilic attack. The substrate is located at 2.570 Å from the cerium and at 4.000 Å from the C5 atom whereas Ce–O distance for the Asp299 is 2.416 Å. This residue behaves as a base and is oriented in a suitable position to

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

accept the proton from the substrate as can be verified by the presence of a hydrogen bond of 1.518 Å.

The TS1 describes the concerted nucleophilic attack and the proton transfer as revealed by the analysis of the imaginary frequency (281*i* cm^{-1}). This value implies a major contribution from the stretching of O_{sub}–C5 bond with respect to that of proton transfer. In fact, the hydrogen lies at 1.037 Å from the Asp299 and at 1.541 Å from the methanol. The distance related to the nucleophilic process is found to be 1.894 Å.

C5 atom assumes a tetrahedral-like structure as confirmation of the deviation from planarity of this moiety (see Figure 3). In this step, the formation of the Ce–N coordination (2.716 Å) compensates the loss of the coordination of the metal with the substrate and this ensures the CN = 9 on all the stationary points isolated on the PES. The barrier required for this step is relatively low (6.5 kcal mol^{-1}) and the formed intermediate (INT1) lies at 3.67 kcal mol^{-1} above the ES species. The INT1 topology well accounts for the concerted processes as evidenced by the newly formed bond between the substrate oxygen and C5 of the PQQ (1.480 Å). Here the H1 of the methyl group of substrate and O4 of PQQ, involved in the next step, are included in a distance of 2.930 Å. In the next step, the transfer of hydrogen from the methyl group to the O4 occurs simultaneously with the breaking of the O–C5 bond realizing the partial PQQ reduction. In fact, looking the TS2_A structure, the migrating H lies at 1.300 and 1.346 Å from the donor and acceptor atoms, respectively. The cleavage of the O–C5 bond is testified by its elongation (2.371 vs. 1.480 Å in INT1) as also indicated by the obtained imaginary frequency (989*i* cm^{-1}). The value of the barrier (19.4 kcal mol^{-1}) shows that this process is the rate-determining step. INT2_A exhibits the departure of the formaldehyde that is moved outside of the reaction (site), although it is retained in the catalytic cavity by the H-bond with Ser169 (2.020 Å), suggesting that the reaction product is already formed. The reaction proceeds until to the accomplishment of the reduction of the cofactor from PQQH to $PQQH_2$ (INT2_A→TS3_A→EP) and occurs below the ES energy. To do this, the protonated Asp299 donates its proton to C5–O hydroxylate group (PQQH) giving back its catalytically active form for another cycle. The barrier for this hydrogen transfer is 5.40 kcal mol^{-1} (TS3_A) with respect to INT2_A even if below the ES energy.

The final species namely EP lies at 8.00 kcal mol^{-1} below the ES proposing a favorable thermodynamics of the process. In the optimized EP the OH groups of the $PQQH_2$ are involved in two strong H-bonds with the Asp299 (1.505 and 1.942 Å) giving rise to a shortening of the three coordination bonds with the metal ion.

Concerning the previous computation on calcium-containing MDH and the same reaction mechanism,[8] a comparison can be done taking into account that in the two active sites some differences (the coordination number and amino acid residues composition) and similarities (both ions are hard in nature) exist. In any case, the topology of the active site for the two enzymes is very similar as shown by the superposition of the two structure (see Figure 4).

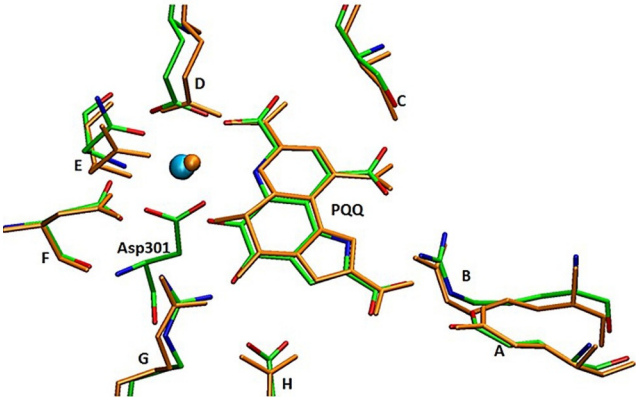


Figure 4. Superposition of Ca-MDH active site (orange) and Ce-MDH; the considered residues are, respectively: A) Glu55–Glu55, B) Arg109–Arg110, C) Ser168–Ser169, D) Glu171–Glu172, E) Asn255–Asn256, F) Asp297–Asp299, G) Arg324–Arg326, H) Asn387–Asp388.

From this Figure, we show the presence of negatively charged residues (Asp/Glu) in the binding cavity for both metal ions that accents the affinity toward cerium, in confirmation of its tendency to substitute Ca^{2+} in protein binding sites. To complete the coordination environment the cerium presents an additional residue (Asp301) enhancing the limited flexibility in the binding cavity typical of the Ca^{2+} containing enzymes.[51] The comparison between the calculated PESs for Ca-MDH and Ce-MDH is difficult since the computational protocols are different (i.e., in the former, the dispersion contributions are neglected). In any case, we note that the two PESs have similar behavior and both indicate TS2 energy as the rate-determining step (19.0 and 21.4 kcal mol^{-1} above the ES for Ca-MDH and Ce-MDH, respectively). The NBO analysis for the ES complexes, depicted in Figure 5, reveals a more negative charge on the substrate oxygen in the case of Ce-MDH due to the greater Lewis acidity of the cerium cation.

This charge distribution makes the O–H substrate bond more polarized, and consequently facilitates the H-transfer and the nucleophilic attack on PQQ that occurs in the first step of the reaction. This can explain the different TS1 energies found in the two systems. In the second part of the reaction, in which the PQQ cofactor plays a principal role, the HOMO composition (see Figure 6) of INT1 can give some indication that

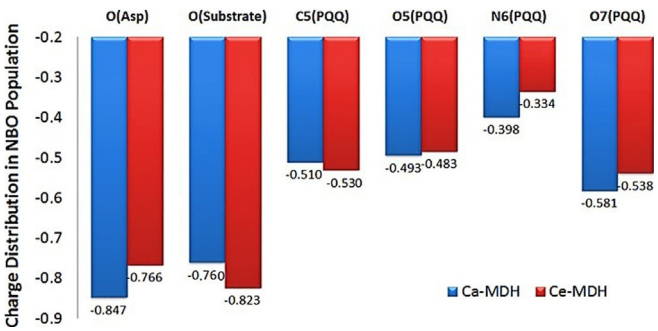


Figure 5. NBO charges (|e|) of selected atoms for ES-complexes.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

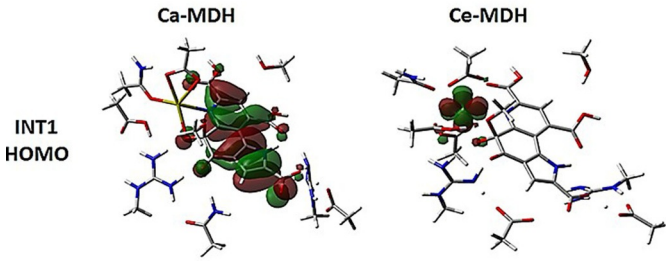


Figure 6. Highest occupied molecular orbital (HOMO) composition of INT1 intermediate, for Ca-MDH and Ce-MDH.

explains the different energetic behavior. In Ce-MDH, the HOMO is mainly due to the d orbitals whereas it is essentially located on the PQQ in the calcium counterpart with a consequent destabilization of the INT1 energy. Starting from this point, all the Ca-MDH energies lie above the corresponding Ce-MDH one.

Conclusions

The reaction mechanism of methanol oxidation to formaldehyde by cerium-containing MDH enzyme has been studied at the DFT level and in the framework of a cluster model, computing the potential energy surfaces for two plausible reaction pathways. Results show that:

1) The preferred mechanism is that in which the addition, elimination and protonation events sequentially occurs (path A)
2) The rate-determining step corresponds to the formaldehyde formation (elimination step) and requires 19.4 kcal mol^{-1} in protein environment
3) Cerium (III) cation, although reported as a nonbiogenic metal species is able to efficiently catalyze the considered reaction due to its Lewis acid character
4) Comparison with Ca-MDH, that follows the same reaction mechanism, shows potential energy surfaces with some similarities and differences. In particular we underline as, in the first part of the reaction, the presence of Ce^{III} center favors the formation of the nucleophilic agent and the produced intermediate INT1_A is thermodynamically stabilized in Ce-MDH. This fact can explain the different PES's behaviors in the addition and elimination steps.

Acknowledgements

We thank Università della Calabria for the financial support.

Conflict of interest

The authors declare no conflict of interest.

Keywords: catalytic mechanism · cerium · density functional calculation · methanol dehydrogenase · potential energy surface

[1] C. Anthony, *Subcell. Biochem.* **2000**, *35*, 73 – 117.
[2] M. Ghosh, C. Anthony, K. Harlos, M. G. Goodwin, C. Blake, *Structure* **1995**, *3*, 177 – 187.
[3] Z. Xia, W. Dai, Y. Zhang, S. A. White, G. D. Boyd, F. S. Mathews, *J. Mol. Biol.* **1996**, *259*, 480 – 501.
[4] Z. X. Xia, Y. N. He, W. W. Dai, S. A. White, G. D. Boyd, F. S. Mathews, *Biochemistry* **1999**, *38*, 1214 – 1220.
[5] Y. J. Zheng, T. C. Bruice, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11881 – 11886.
[6] Y. J. Zheng, Z. Xia, Z. Chen, F. S. Mathews, T. C. Bruice, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 432 – 434.
[7] N. B. Idupulapati, D. Mainardi, *J. Phys. Chem. A* **2010**, *114*, 1887 – 1896.
[8] M. Leopoldini, N. Russo, M. Toscano, *Chem. Eur. J.* **2007**, *13*, 2109 – 2117.
[9] P. Hothi, M. J. Sutcliffe, N. S. Scrutton, *Biochem. J.* **2005**, *388*, 123 – 133.
[10] X. Zhang, S. Y. Reddy, T. C. Bruice, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 745 – 749.
[11] S. Y. Reddy, T. Bruice, *Protein Sci.* **2004**, *13*, 1965 – 1978.
[12] N. B. Idupulapati, D. S. Minardi, *THEOCHEM* **2009**, *901*, 72 – 80.
[13] A. Pol, T. R. M. Barends, A. Dietl, A. F. Khadem, J. Eygensteyn, M. S. M. Jetten, H. J. M. Op den Camp, *Enviromental Microbiol.* **2014**, *16*, 255 – 264.
[14] A. M. Panichev, *Achiev. Life Sci.* **2015**, *9*, 95 – 103.
[15] M. Murugesu, E. J. Schelter, *Inorg. Chem.* **2016**, *55*, 9951 – 9953.
[16] K. Djinovic-Carugo, O. Carugo, *J. Inorg. Biochem.* **2015**, *143*, 69 – 76.
[17] R. B. Martin, F. S. Richardson, *Q. Rev. Biophys.* **1979**, *12*, 181 – 209.
[18] C. G. dos Remedios, *Cell Calcium* **1981**, *2*, 29 – 51.
[19] R. B. Mikkelson, *Lanthanides as Calcium Probes, in Biological Membranes, Vol. III* (Eds.: D. Chapman, D. F. Wallach), Academic Press, New York, **1976**, pp. 153 – 190.
[20] J. A. Bogart, A. J. Lewis, E. J. Schelter, *Chem. Eur. J.* **2015**, *21*, 1743 – 1748.
[21] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, **2013**, Gaussian, Inc., Wallingford CT.
[22] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.
[23] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785 – 789.
[24] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456 – 1465.
[25] D. Andrae, H. Haußermann, M. Dolg, H. Stoll, H. Preuß, *Theor. Chim. Acta* **1990**, *77*, 123 – 141.
[26] M. R. A. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601 – 3658, and references therein.
[27] A. J. M. Ribeiro, M. E. Alberto, M. J. Ramos, P. A. Fernandes, N. Russo, *Chem. Eur. J.* **2013**, *19*, 14081 – 14089.
[28] P. Piazzetta, T, Marino, N. Russo, *Phys. Chem. Chem. Phys.* **2014**, *16*, 16671 – 16676.
[29] A. J. M. Ribeiro, D. Santos-Martins, N. Russo, M. J. Ramos, P. A. Fernandes, *ACS Catal.* **2015**, *5*, 5617 – 5626.
[30] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* **2003**, *24*, 669 – 681.
[31] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995 – 2001.
[32] S. Chen, W. H. Fang, F. Himo, *J. Inorg. Biochem.* **2009**, *103*, 274 – 281.
[33] R. Z. Liao, F. Himo, J. G. Yu, R. Z. Liu, *J. Inorg. Biochem.* **2010**, *104*, 37 – 46.
[34] S. F. Sousa, P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **2009**, *15*, 4243 – 4247.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[35] M. E. Alberto, T. Marino, M. J. Ramos, N. Russo, *J. Chem. Theory Comput.* **2010**, *6*, 2424 – 2433.
[36] O. Amata, T. Marino, N. Russo, M. Toscano, *Phys. Chem. Chem. Phys.* **2011**, *13*, 3468 – 3477.
[37] P. E. M. Siegbahn, L. A. Eriksson, F. Himo, M. Pavlov, *J. Phys. Chem. B* **1998**, *102*, 10622 – 10629.
[38] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley, New York, **1991**.
[39] E. D. Glendening, A. E. Reed, J. E. Carpenter, F. Weinhold, NBO, version 3.1.
[40] P. Piazzetta, T. Marino, N. Russo, *Inorg. Chem.* **2014**, *53*, 3488 – 3493.
[41] O. Amata, T. Marino, N. Russo, M. Toscano, *J. Am. Chem. Soc.* **2011**, *133*, 17824 – 17831.
[42] R. Z. Liao, J. G. Yu, F. Himo, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22523 – 22527.
[43] M. J. Ramos, P. A. Fernandes, *Acc. Chem. Res.* **2008**, *41*, 689 – 698.
[44] T. Marino, N. Russo, M. Toscano, *Chem. Eur. J.* **2013**, *19*, 2185 – 2192.
[45] C. Anthony, *Biochem. J.* **1996**, *320*, 697 – 711.
[46] J. J. Frank, S. H. van Krimpen, P. E. Verwiel, J. A. Jongenjan, A. C. Malder, J. A. Duine, *Eur. J. Biochem.* **1989**, *184*, 187 – 195.
[47] A. J. Olsthoorn, J. A. Duine, *Biochemistry* **1998**, *37*, 13854 – 13861.
[48] J. J. Frank, M. Dijkstra, J. A. Duine, C. Balny, *Eur. J. Biochem.* **1988**, *174*, 331 – 358.
[49] L. Noodleman, T. Lovell, W.-G. Han, J. Li, F. Himo, *Chem. Rev.* **2004**, *104*, 459 – 508.
[50] F. Chu, M. E. Lidstrom, *J. Bacteriol.* **2016**, *198*, 1317 – 1325.
[51] T. Dudev, C. Lim, *Chem. Rev.* **2014**, *114*, 538 – 556.

Manuscript received: January 25, 2017
Accepted manuscript online: May 10, 2017
Version of record online: June 7, 2017

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

CHEMISTRY
A European Journal

Supporting Information

How Can Methanol Dehydrogenase from *Methylacidiphilum fumariolicum* Work with the Alien Ce^{III} Ion in the Active Center? A Theoretical Study

Mario Prejanò, Tiziana Marino,* and Nino Russo[a]

chem_201700381_sm_miscellaneous_information.pdf

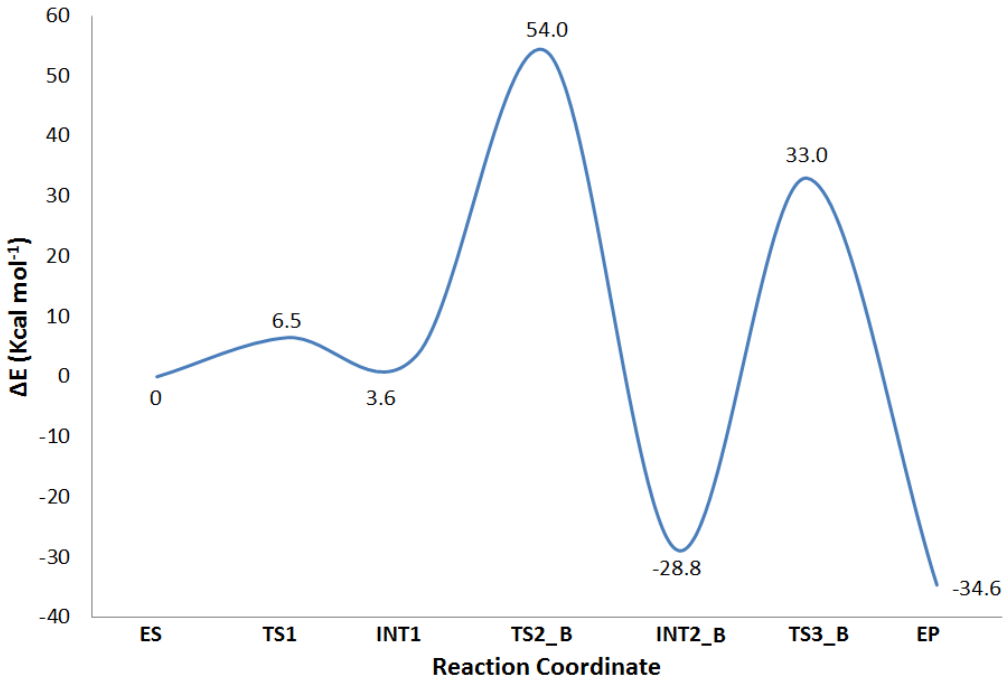


Figure S1. The Potential Energy Surface (PES) for the addition-retro-ene -reaction (mechanism B), calculated with C-PCM(ε=4.0)/B3LYP-D3/6-311+G(2d,2p)|SDD.

Table S1. ES coordinates

C	-11.09504800	-1.67161400	-0.84394800
C	-9.76163700	-0.93749000	-0.78788600
O	-8.75766000	-1.50614100	-1.29891300
O	-9.75353300	0.21464000	-0.23702900
C	-8.79994300	1.54069400	3.43663800
N	-8.57074000	1.29717500	2.02144200
C	-7.41195100	1.51494700	1.39038900
N	-6.31061300	1.89387600	2.08800000
N	-7.37798300	1.43120400	0.05883700
C	-3.40364400	8.22961100	-0.52101800
O	-3.38329200	6.79791400	-0.42368600
C	4.24494300	1.80774500	-4.18310900
C	4.13698700	1.27448900	-2.77156000
O	5.17954500	1.27682000	-2.03116700
O	3.03286500	0.83254000	-2.33672500
C	7.55701000	2.57800200	1.98331500
C	6.66660600	2.33861900	0.78764600
O	5.44781200	2.14505500	0.95348100
N	7.22765700	2.32623600	-0.43327100
C	8.25558500	-1.36328500	1.08528700
C	6.95683100	-1.38161900	0.29688500
O	6.06264500	-0.53868200	0.65558000
O	6.82441700	-2.20018800	-0.64892400
C	2.18134600	-0.41682500	4.02034200
C	2.76698400	-0.13957700	2.65386100
O	3.28858000	-1.10248800	1.98853400
O	2.71908300	1.02756500	2.16912700
C	1.24464400	-3.74872000	2.94495400
N	2.11764500	-3.61230900	1.78068700
C	1.87342100	-4.19718500	0.60730600
N	0.75072100	-4.91266000	0.43937200
N	2.78511700	-4.10823400	-0.39950400
C	-2.64919500	-5.07392300	1.28449700
C	-2.36643400	-3.55628100	1.23816000
O	-1.12689200	-3.23690500	1.03967000
O	-3.29810000	-2.74423300	1.34976600
Ce	3.91946400	0.36902300	0.00872200
N	-3.64129100	-0.49514600	-1.01462900
C	-4.09513000	-1.67049500	-1.53386300
C	-5.54374500	-1.97166000	-1.67723000
O	-5.92198800	-3.02674100	-2.15224000
C	-2.99764400	-2.47297400	-1.83319100
C	-1.85495100	-1.75619400	-1.44279100
C	-2.27915400	-0.49806200	-0.93081000
C	-0.46722300	-2.17483100	-1.43871100
O	-0.00698100	-3.19352400	-1.93842200
C	0.49019300	-1.16058200	-0.77901800
O	1.66234300	-1.43947700	-0.60391300
C	-0.00926400	0.22404200	-0.45883000
N	0.96636000	1.07491100	-0.20208200
C	0.66664100	2.36068100	0.03228400
C	1.83296900	3.26280800	0.17833100
O	2.98272900	2.93295800	-0.05393900
C	-0.64518500	2.81370000	0.04923900
C	-1.70901400	1.93598100	-0.21309000
C	-3.06562900	2.54982100	-0.17105900
O	-4.12705200	1.96708000	0.04148800
C	-1.40986200	0.56497800	-0.51136800
H	-10.98273900	-2.65156200	-1.30982600
H	-11.49367900	-1.78239000	0.17005600
H	-8.29628800	0.80012800	4.07250400
H	-9.87211500	1.47127100	3.62436000
H	-9.30011800	0.87423300	1.43202100
H	-5.43775800	1.94437700	1.57441900
H	-6.23960700	1.63043200	3.05925500

H	-6.49688400	1.21570900	-0.38697800
H	-8.23619000	0.99465500	-0.35102200
H	-2.67039700	8.69199100	0.15280700
H	-3.21971900	8.57023200	-1.54847700
H	-2.51244900	6.47566600	-0.68842100
H	0.93624000	-4.79176500	3.05160900
H	0.34605300	-3.13291600	2.84621300
H	2.74153700	-2.79843300	1.80079100
H	0.65635100	-5.32233500	-0.48160200
H	-0.12608900	-4.36608600	0.75115700
H	3.52524800	-3.41776200	-0.28833900
H	2.33836100	-3.99887100	-1.30372200
H	-3.67474000	-5.26391200	1.60640900
H	-1.94377700	-5.58080100	1.95118500
H	3.29995200	1.69394400	-4.71623200
H	5.04279700	1.27754500	-4.71238000
H	8.55090500	2.95117700	1.72145300
H	7.65706200	1.62272500	2.50930100
H	6.64820200	2.04382200	-1.22878500
H	8.72547700	-0.37906200	0.98135000
H	8.93935700	-2.13666900	0.73283200
H	1.17874300	-0.84025500	3.89126800
H	2.09972500	0.50743500	4.59468100
H	-4.23122200	0.23423800	-0.63529100
H	-3.04054300	-3.47185800	-2.23866800
H	-0.84082200	3.85608800	0.25691100
H	-11.81342600	-1.06850900	-1.40899200
H	-2.51143200	-5.49143000	0.27951000
H	1.81623000	-3.45411600	3.82704000
H	2.79031200	-1.15075100	4.55377900
H	7.06359800	3.27975900	2.65943000
H	8.03752100	-1.50988500	2.14796300
H	4.52584300	2.86568200	-4.14677400
H	-4.40441000	8.54751700	-0.22115000
H	-8.46589000	2.54512400	3.71883900
O	-3.02594100	3.87881400	-0.36525200
H	-3.91107700	4.27465800	-0.27152400
O	-6.31037900	-0.97403100	-1.24115400
H	-7.34213200	-1.19865000	-1.27947900
O	1.50784000	4.51848100	0.54855800
H	2.33841200	5.01876900	0.61826100
H	8.22561200	2.40875700	-0.54591600
C	4.04816500	-2.10266700	-2.64597300
H	2.99360300	-1.85742900	-2.77785000
H	4.64218000	-1.50983100	-3.35132700
H	4.21251300	-3.16779900	-2.85276100
O	4.40776300	-1.79494400	-1.29711300
H	5.39816500	-2.01043900	-1.13366400

Table S2. TS1 coordinates

C	10.60812200	-0.54514700	0.30114600
C	9.19565000	-0.01584600	0.27749400
O	8.54792900	-0.25049900	1.38348300
O	8.72930400	0.55312800	-0.72164400
C	6.87249000	2.03981100	-3.34209800
N	6.56568700	1.03455800	-2.31663900
C	5.61906300	0.09178500	-2.39440600
N	4.88808800	-0.04748600	-3.53659800
N	5.35222800	-0.68621200	-1.34640500
C	0.81946900	7.14353900	1.83837700
O	0.47087500	6.58933500	0.57085400
C	-5.51120700	3.41195200	1.86558700
C	-4.84557200	2.28240700	1.11116700
O	-5.54973500	1.58132100	0.30144600
O	-3.62178600	2.02855200	1.28458800
C	-7.32629600	-0.83212600	-3.42273700
C	-6.67743600	-0.12431400	-2.25563600

O	-5.50659500	-0.41415200	-1.93725700
N	-7.38275300	0.80286300	-1.59143300
C	-5.08288100	-2.06934000	3.73588900
C	-4.20555400	-1.15605100	2.91925000
O	-4.37710300	-1.04418000	1.69166900
O	-3.28008900	-0.55857900	3.61112100
C	-1.21199300	-2.64001500	-2.88061200
C	-2.21716100	-1.76686800	-2.16872400
O	-3.08267500	-2.28555200	-1.38105800
O	-2.15405500	-0.50409400	-2.32262900
C	-1.31537400	-5.73648800	-1.32978900
N	-1.97396100	-4.75690000	-0.46856000
C	-1.45035500	-4.34782500	0.69926300
N	-0.21946600	-4.72742000	1.05680500
N	-2.21763900	-3.59939000	1.53019800
C	2.99656600	-5.91350400	-0.52481300
C	2.70038100	-4.44349000	-0.93822200
O	1.46480600	-4.10017400	-0.85948100
O	3.63737100	-3.70981600	-1.33098000
Ce	-3.57920400	0.07862800	-0.39573600
N	3.41641100	0.70825200	1.17423000
C	3.93775600	-0.24982300	1.99559700
C	5.41993000	-0.39292500	2.16547700
O	5.85402200	-1.00966100	3.13519000
C	2.89309500	-1.02044100	2.49441900
C	1.71133900	-0.51261200	1.92927500
C	2.05454300	0.59687100	1.10397900
C	0.37481500	-1.07062400	1.98413100
O	0.07536500	-2.11099000	2.56589100
C	-0.66835800	-0.29851200	1.19148500
O	-1.70071000	-0.91153900	0.79413700
C	-0.22793500	0.88699500	0.36669300
N	-1.21143500	1.40777900	-0.33563900
C	-1.00862700	2.51634800	-1.06953400
C	-2.19925700	2.95805100	-1.82498600
O	-3.27677600	2.37192600	-1.79286400
C	0.22328700	3.15514900	-1.08696400
C	1.30947600	2.59092800	-0.39540200
C	2.61018600	3.29829100	-0.52794100
O	3.72384100	2.80483800	-0.51801800
C	1.11856000	1.38662700	0.35154100
H	11.08852000	-0.30405600	1.25286600
H	10.56985800	-1.63733000	0.22266900
H	7.35276300	1.59813900	-4.22551300
H	7.56690900	2.75391900	-2.89911800
H	7.25927800	0.91845600	-1.55719000
H	4.21183100	-0.79889500	-3.56981300
H	5.30690900	0.20792200	-4.41819900
H	4.77072100	-1.52184100	-1.42596000
H	5.69693500	-0.41915300	-0.40615500
H	0.83416800	6.38310900	2.63242900
H	1.79523700	7.65071900	1.81511100
H	1.12524100	5.92556000	0.31392800
H	-1.04367500	-6.61403700	-0.73679500
H	-0.40699500	-5.33703900	-1.79129800
H	-2.62497000	-4.09270700	-0.89398600
H	0.09403800	-4.34812400	1.94286300
H	0.51840700	-4.66110600	0.28310200
H	-3.05556700	-3.21381200	1.11168600
H	-1.68431800	-2.89884600	2.04857900
H	4.07145800	-6.09320600	-0.45354100
H	2.56767300	-6.58446200	-1.27818000
H	-6.22704300	2.98999300	2.57939700
H	-6.07266400	4.04352700	1.17080800
H	-8.35938700	-0.52295700	-3.60082000
H	-7.29411000	-1.90912100	-3.23607100
H	-6.93158100	1.27802400	-0.79992300
H	-5.22562700	-1.67229400	4.74271400
H	-4.57208000	-3.03555000	3.82418100
H	-0.38196900	-2.87171400	-2.19801100

H	-0.80859800	-2.11546200	-3.74849100
H	3.98262000	1.34994100	0.62802000
H	2.99303800	-1.86961000	3.15237100
H	0.34786100	4.07153400	-1.64864100
H	11.18239500	-0.14379300	-0.53493400
H	2.51667500	-6.14312400	0.43257200
H	-2.02642600	-6.03961800	-2.10141000
H	-1.66628900	-3.58400600	-3.19014300
H	-6.72937800	-0.63884600	-4.31875700
H	-6.03793300	-2.22226800	3.23194400
H	-4.77069200	4.00600300	2.40329300
H	0.05445700	7.88403800	2.08570600
H	5.96370300	2.56905100	-3.64682600
O	2.45301200	4.63951700	-0.70907100
H	3.33934500	5.01468600	-0.84830400
O	6.09538100	0.17477300	1.20677000
H	7.52269300	-0.02281300	1.33247400
O	-2.01308200	4.06216600	-2.56425800
H	-2.84922500	4.27092600	-3.01409500
H	-8.33059600	1.02345100	-1.85235500
C	-0.88404900	1.94933700	3.00148100
H	-1.38805200	2.59545900	2.27045100
H	-1.24818000	2.21685100	4.00320200
H	0.20054800	2.14091100	2.98467400
O	-1.16236300	0.58311800	2.79052900
H	-2.53574900	-0.03750900	3.11081500

Table S3. INT1 coordinates

C	10.47060900	0.09000500	1.59447800
C	8.98067700	0.08626500	1.34175900
O	8.28960300	-0.43693200	2.30942800
O	8.47807900	0.54014400	0.30081800
C	7.61449100	2.36654400	-2.94092100
N	7.22128500	1.29659600	-2.00684000
C	6.27780000	0.36190200	-2.20858500
N	5.64745500	0.29479700	-3.41740200
N	5.91109100	-0.47338400	-1.24039300
C	0.36853900	7.43110100	0.43120900
O	1.42658700	6.86953000	-0.34163300
C	-5.40273600	3.10668400	2.73402400
C	-4.78804900	2.16107800	1.72242600
O	-5.26864300	0.97694800	1.62307100
O	-3.81186200	2.53077500	1.01585200
C	-7.85121600	-0.07926000	-2.35747100
C	-6.93549200	0.25232000	-1.20387100
O	-5.73276000	0.50130300	-1.42951400
N	-7.43535000	0.26706700	0.04000500
C	-5.53899000	-2.97972300	2.79614500
C	-4.50498400	-2.02605600	2.26720400
O	-4.46074200	-1.67950300	1.08306900
O	-3.65160700	-1.62486500	3.18692600
C	-2.06020100	-2.09950600	-3.63474200
C	-2.71569400	-1.23685900	-2.57925000
O	-3.58652100	-1.74116600	-1.79631900
O	-2.35033200	-0.02406300	-2.45644900
C	-0.58079700	-5.64601600	-2.49738400
N	-1.44208000	-5.23470200	-1.39242800
C	-1.33605000	-4.08109000	-0.70238300
N	-0.24403700	-3.34231800	-0.72233600
N	-2.41019100	-3.71799400	0.06347400
C	3.42689900	-5.61371300	-0.70582200
C	3.26993000	-4.19060100	-1.31542000
O	2.09704900	-3.91942500	-1.76337500
O	4.25800700	-3.42062100	-1.35695400
Ce	-3.52791300	0.29865500	-0.28499500
N	3.28627600	0.14753200	1.25634800
C	3.66767200	-0.89344100	2.05456500

C	5.11798500	-1.18278400	2.30574200
O	5.41151100	-2.05272700	3.12448400
C	2.52359700	-1.50532200	2.55386200
C	1.42342900	-0.80918000	2.01873900
C	1.92662300	0.24657900	1.20110500
C	0.01139100	-1.04119400	2.23716800
O	-0.46540200	-1.82785400	3.04766400
C	-0.93982300	-0.25244300	1.31107000
O	-1.55080900	-1.04992100	0.42123100
C	-0.29398500	0.91972400	0.54330400
N	-1.19526800	1.57187100	-0.15773700
C	-0.84583400	2.61446300	-0.92882900
C	-1.96531000	3.18246800	-1.71906600
O	-3.08850100	2.70576400	-1.73540900
C	0.46585700	3.05892200	-0.96746600
C	1.46490400	2.35213100	-0.27197200
C	2.84878700	2.88854600	-0.38749800
O	3.89482400	2.26818900	-0.32161400
C	1.10933400	1.19137600	0.48433500
H	10.68706000	0.52386800	2.56824200
H	10.82415500	-0.95300300	1.61868500
H	8.20498700	1.98310100	-3.78264800
H	8.22561900	3.07711100	-2.38535800
H	7.77694200	1.17331800	-1.15132200
H	4.97678400	-0.45094800	-3.55218400
H	6.14860500	0.58497900	-4.24360000
H	5.31910800	-1.28855100	-1.41909500
H	6.19616700	-0.33468700	-0.26230100
H	-0.59408100	6.94153600	0.24062800
H	0.57981800	7.39595100	1.51143500
H	1.55466000	5.94409200	-0.09590600
H	-0.60798600	-6.73600100	-2.57063700
H	0.45293900	-5.33477900	-2.31972800
H	-2.28067500	-5.76778300	-1.21590600
H	-0.35267400	-2.41349200	-0.31748300
H	0.70018600	-3.63461200	-1.11248700
H	-3.29364900	-3.71841100	-0.43667800
H	-2.27617200	-2.81068200	0.51471200
H	4.41902100	-5.75139800	-0.27116800
H	3.26245000	-6.35721500	-1.49412300
H	-4.96093900	4.10140800	2.65672500
H	-5.23760200	2.70875900	3.74135500
H	-8.88776100	-0.25255000	-2.05641400
H	-7.46087200	-0.97279800	-2.85344600
H	-6.79989500	0.50574800	0.81514300
H	-6.02686400	-2.54529500	3.67381700
H	-5.04245500	-3.90061100	3.11782100
H	-1.06397000	-2.38048200	-3.27484900
H	-1.93290900	-1.53365100	-4.56053400
H	3.94179000	0.73637000	0.75141800
H	2.50269500	-2.36093100	3.21113500
H	0.71829400	3.93184900	-1.55421900
H	10.99337200	0.62821800	0.80341200
H	2.65979400	-5.76845600	0.06070900
H	-0.92286600	-5.21338400	-3.44545900
H	-2.64348300	-3.00539300	-3.81303800
H	-7.81190100	0.74011200	-3.07960900
H	-6.27132300	-3.20189400	2.02050300
H	-6.48508200	3.16167600	2.58270700
H	0.28473000	8.48150500	0.13557000
H	6.72998500	2.88575500	-3.32487600
O	2.84619200	4.22822700	-0.60841000
H	3.76832800	4.52120200	-0.69893100
O	5.91944400	-0.42229800	1.62271800
H	7.25238500	-0.46944800	2.05458600
O	-1.62382800	4.27531600	-2.43011600
H	-2.40949200	4.56000500	-2.92584600
H	-8.40718000	0.06503900	0.21476200
C	-1.72699900	1.31099800	3.09674900
H	-1.46837000	2.24783200	2.59455200

H	-2.63066200	1.46180000	3.68844800
H	-0.91459200	0.98008500	3.75297400
O	-2.04172300	0.29683700	2.13156100
H	-2.95444700	-1.02587000	2.80390200

Table S4. TS2_A coordinates

C	-10.63311500	-0.80468400	-2.26377400
C	-9.31377000	-0.40851400	-1.63162200
O	-8.28493100	-0.85324600	-2.29606200
O	-9.26337900	0.25705800	-0.58553500
C	-8.59271600	2.20500300	2.28475600
N	-7.83610000	1.14187900	1.60591300
C	-6.81856400	0.44690000	2.13353900
N	-6.47424200	0.65174000	3.43864600
N	-6.11670200	-0.40935900	1.39832300
C	-1.97262000	8.16469500	-0.90597600
O	-2.32830800	6.94878500	-0.23977700
C	2.75470100	2.74556800	-3.13680100
C	3.34681900	2.06489700	-1.93191400
O	4.05878900	1.00705100	-2.08557100
O	3.09325200	2.50154900	-0.76851400
C	8.44279000	1.91964100	-0.24303100
C	7.04395800	1.58099400	-0.69984800
O	6.12522800	1.49989400	0.14089300
N	6.82134500	1.36898500	-2.00532600
C	5.85976300	-2.68280900	-2.79113100
C	4.79677400	-2.00575400	-1.97100900
O	5.04389700	-1.44958600	-0.89146300
O	3.59586100	-2.07007600	-2.49610900
C	5.04720000	-2.45389100	3.62030600
C	4.60975400	-1.48165300	2.54605900
O	3.54364900	-1.73185800	1.88040600
O	5.28057800	-0.43096600	2.32866700
C	0.77924200	-4.35520400	3.11488100
N	1.64677500	-3.87441500	2.04315600
C	1.33353300	-3.98501600	0.74002300
N	0.11977600	-4.37765500	0.36945200
N	2.31170600	-3.74424700	-0.17703600
C	-3.00590400	-5.10300000	1.02097600
C	-2.95306800	-3.61674600	1.45391000
O	-1.76597400	-3.11119700	1.49991300
O	-4.00377700	-2.99894600	1.71832600
Ce	3.89238400	0.42688700	0.42168000
N	-3.29134400	-0.01259100	-0.72755800
C	-3.69903700	-1.05134700	-1.50306800
C	-5.13551000	-1.27618800	-1.88307600
O	-5.38348800	-2.16013200	-2.70262900
C	-2.57877100	-1.79932400	-1.87980700
C	-1.46434100	-1.18523400	-1.29955100
C	-1.91965100	-0.04455100	-0.56952800
C	-0.06575100	-1.51106500	-1.40629300
O	0.39210300	-2.42648700	-2.19776800
C	0.86632800	-0.68053100	-0.70779200
O	2.13665300	-0.88474000	-0.73158700
C	0.35114800	0.41389000	0.13206400
N	1.28503700	1.01519200	0.85065500
C	0.98988900	2.11664600	1.56119500
C	2.13264700	2.67158800	2.29214900
O	3.26339100	2.17551400	2.26552300
C	-0.27405900	2.70564900	1.48927700
C	-1.29464400	2.08966800	0.76478200
C	-2.58055800	2.83848700	0.66650600
O	-3.69846400	2.36761900	0.54334600
C	-1.04586100	0.82796300	0.11599100
H	-10.63732800	-0.52388800	-3.32080700
H	-10.73542600	-1.89375200	-2.21607500
H	-9.23442100	1.80989700	3.08377300

H	-9.23161900	2.67413300	1.53642400
H	-8.21399900	0.80181500	0.70451600
H	-5.71880300	0.10105000	3.82025300
H	-7.18088800	0.94989600	4.09321900
H	-5.45521800	-1.07076800	1.80466100
H	-6.16397700	-0.41583500	0.35605900
H	-1.07033400	8.61879000	-0.47356400
H	-1.81780400	8.01094300	-1.98269900
H	-1.60924700	6.30827000	-0.32107700
H	1.36275900	-4.37926900	4.03809100
H	0.44166100	-5.36877400	2.88122400
H	2.39916200	-3.21706800	2.25575200
H	-0.00802600	-4.43695700	-0.63451900
H	-0.74976300	-3.90458600	0.91511900
H	3.10966100	-3.22647900	0.17480700
H	1.97149900	-3.40324900	-1.07078800
H	-2.28472300	-5.69485300	1.59281200
H	-2.73199500	-5.17915000	-0.03774600
H	2.63053600	3.81505400	-2.95400100
H	1.76736500	2.29390000	-3.29304800
H	9.16734800	1.97360700	-1.05947000
H	8.75984300	1.16167500	0.47890800
H	5.85458600	1.18183600	-2.30012900
H	5.90790700	-2.20737700	-3.77593900
H	5.58821000	-3.73068900	-2.95028900
H	4.91914900	-3.48400100	3.27582200
H	4.40905500	-2.31400600	4.50068300
H	-3.90467300	0.68168600	-0.31546700
H	-2.60508100	-2.67901700	-2.50307000
H	-0.45258900	3.65131300	1.98437900
H	-11.46424400	-0.33172500	-1.73936000
H	-4.01199700	-5.50650700	1.14971900
H	-0.10340800	-3.72029300	3.24678600
H	6.08426000	-2.27238600	3.90638100
H	8.41332200	2.87849900	0.28211400
H	6.82423800	-2.61250800	-2.28895600
H	3.36141600	2.56557600	-4.02705800
H	-2.80851600	8.85508900	-0.76982600
H	-7.91506500	2.95899100	2.69814000
O	-2.37844800	4.17756300	0.73145600
H	-3.23489900	4.63357400	0.65476700
O	-5.99323400	-0.47667500	-1.31230300
H	-7.34648100	-0.67486600	-1.86082600
O	1.86959300	3.77551200	3.00881200
H	2.70052900	4.06117100	3.42446300
H	7.56115500	1.45642000	-2.68374100
C	0.92352700	-0.52531600	-3.79740900
H	0.76810000	-1.70184200	-3.26764900
H	0.21460200	-0.64861900	-4.63720600
H	1.99091600	-0.54285500	-4.08011400
O	0.56688100	0.29400200	-2.84889400
H	2.92144400	-1.58771300	-1.90419900

Table S5. TS2_B coordinates

C	-10.66908200	-0.28118400	-1.14462300
C	-9.21301200	-0.04784200	-0.80944400
O	-8.40712800	-0.66979000	-1.62239300
O	-8.86538900	0.65853200	0.14971300
C	-7.47189200	2.49749900	2.85122200
N	-7.00789500	1.33866200	2.07259400
C	-6.02737800	0.49675700	2.42678500
N	-5.43295000	0.63889100	3.64846800
N	-5.59788400	-0.44642800	1.59424500
C	-0.85831200	7.51285400	-1.69098600
O	-1.52191000	7.03162900	-0.52348000
C	6.33775100	2.67599600	-2.33372000
C	5.41452500	1.88088700	-1.43606300

O	5.84713700	1.48412000	-0.29815200
O	4.23150600	1.62463800	-1.80068200
C	6.79111400	-0.50867300	3.95112200
C	6.39149300	0.10343300	2.62855300
O	5.20521000	0.02333500	2.25108200
N	7.32993000	0.71984200	1.89573300
C	5.86630500	-2.54232600	-2.82038300
C	4.67281900	-1.91364500	-2.14642600
O	4.76442500	-1.53401600	-0.94445400
O	3.60509400	-1.78637400	-2.84316500
C	0.82996300	-1.77395800	3.01482300
C	1.91439100	-1.09121600	2.21476500
O	2.77964800	-1.78652600	1.57845400
O	1.93240500	0.17882900	2.14239100
C	1.00281300	-5.20554900	2.21637400
N	1.75961700	-4.41495100	1.25174100
C	1.35376600	-4.20557000	-0.01422700
N	0.14406300	-4.60706600	-0.41585900
N	2.22955900	-3.64281000	-0.87364900
C	-3.16502600	-5.48537900	0.86611500
C	-2.89576700	-4.00832400	1.26762800
O	-1.66476900	-3.64195500	1.18228200
O	-3.84051900	-3.28162200	1.64827900
Ce	3.64178300	0.29691300	0.30198600
N	-3.30357600	0.17154400	-1.11759800
C	-3.76159000	-0.96193200	-1.72400500
C	-5.22727200	-1.26173400	-1.81075200
O	-5.59630100	-2.19663700	-2.51755700
C	-2.66699800	-1.67456700	-2.20741200
C	-1.52255000	-0.92770000	-1.88058400
C	-1.94214600	0.24626600	-1.19260900
C	-0.12998800	-1.24094000	-2.10529000
O	0.28109400	-2.31018100	-2.55934800
C	0.85757700	-0.08420400	-1.81960400
O	2.09403300	-0.54198800	-1.45297100
C	0.34317300	0.93890600	-0.79389800
N	1.29707300	1.58801100	-0.17155200
C	0.99682400	2.65350300	0.60310900
C	2.15191000	3.22309400	1.32009400
O	3.28243600	2.74217200	1.27514400
C	-0.29370300	3.14246000	0.68719100
C	-1.34624400	2.43663400	0.07515800
C	-2.70104000	3.03611600	0.21842000
O	-3.77656700	2.46297600	0.22142600
C	-1.05778700	1.22911000	-0.63050900
H	-10.85408900	-0.01555000	-2.18977200
H	-10.89275100	-1.34747400	-1.03774700
H	-8.02228500	2.19646200	3.75247600
H	-8.14874400	3.06508700	2.21247000
H	-7.57568200	1.06392600	1.25249000
H	-4.72659100	-0.03566300	3.90608700
H	-5.97071500	1.02357100	4.41037200
H	-4.97851300	-1.20430400	1.88494000
H	-5.83387200	-0.42884100	0.58146900
H	0.18102600	7.15900200	-1.75145900
H	-1.38433400	7.22946100	-2.61386800
H	-1.56022800	6.06609800	-0.54176100
H	0.73428500	-6.16566400	1.76734700
H	0.08199900	-4.70485600	2.53352900
H	2.40229000	-3.70167300	1.59899500
H	-0.06587000	-4.39482900	-1.38509900
H	-0.66550500	-4.37423000	0.26473400
H	3.06505900	-3.23616700	-0.47385700
H	1.83521700	-3.12908400	-1.65803300
H	-2.58291800	-6.14933100	1.51501900
H	-2.83118800	-5.65623800	-0.16344700
H	5.91447100	2.77278300	-3.33459400
H	7.32066100	2.19765300	-2.37682000
H	7.84163100	-0.34890800	4.20699100
H	6.58648400	-1.58231800	3.90868400

H	7.05110000	1.12507400	0.99396400
H	6.50479800	-1.73657600	-3.20113200
H	5.55517700	-3.16570600	-3.65987600
H	0.06349000	-2.17700400	2.33801600
H	0.35704700	-1.06123300	3.69234600
H	-3.90127600	0.84474800	-0.64594800
H	-2.71418500	-2.61814000	-2.72890000
H	-0.49393500	4.04589800	1.24841300
H	-11.31257100	0.30201500	-0.48501000
H	-4.22682000	-5.72560800	0.94676000
H	1.64556000	-5.38928600	3.07997700
H	1.24280200	-2.61210900	3.58217000
H	6.15612400	-0.08655700	4.73489000
H	6.44621000	-3.11797900	-2.09652400
H	6.47667500	3.67371500	-1.90238400
H	-0.84867100	8.60333100	-1.62015900
H	-6.62843100	3.13553600	3.13480000
O	-2.63563400	4.38544800	0.37211800
H	-3.54324600	4.71520000	0.48449500
O	-5.96887000	-0.44528800	-1.11553500
H	-7.38500900	-0.57220200	-1.39426400
O	1.87608300	4.32320200	2.03703000
H	2.69586400	4.60442300	2.47688200
H	8.28762900	0.76453200	2.20581800
C	1.63563600	0.25820800	-4.11617200
H	1.44427800	-0.79558300	-4.34903500
H	1.34215600	0.87848600	-4.96527400
H	2.69664500	0.40484200	-3.89723900
O	0.83274600	0.71735800	-3.01865100
H	2.77032300	-1.18431700	-2.22616800

Table S6. INT2_A coordinates

C	-10.72909200	0.15702000	0.98156300
C	-9.26641800	-0.16819000	0.80503600
O	-8.53475300	0.18621300	1.81998300
O	-8.84448000	-0.71460800	-0.22588500
C	-7.17900900	-1.94616900	-3.12743500
N	-6.87624300	-0.93769600	-2.10823900
C	-5.91441100	-0.01202100	-2.17964600
N	-5.19319800	0.12550400	-3.32763100
N	-5.63083400	0.75523000	-1.12986500
C	-0.44273800	-7.01609900	1.16712000
O	-0.22722800	-6.29115900	-0.04556300
C	4.90299200	-3.44438400	2.36154600
C	4.52138100	-2.33742500	1.40201200
O	5.16239500	-1.23126800	1.44837800
O	3.56044700	-2.49848900	0.59765400
C	8.11973800	0.42014300	-2.04291400
C	7.09066400	-0.11781400	-1.07804400
O	5.89248100	-0.18274300	-1.42978300
N	7.48721300	-0.51759700	0.13707800
C	5.71377200	1.72014000	3.59088100
C	4.58887800	1.40239400	2.64786700
O	4.67528700	1.53537800	1.42163900
O	3.49875000	0.95648700	3.24631200
C	2.45512600	3.16688100	-3.13075500
C	2.94455900	2.04266600	-2.24503700
O	3.77766100	2.28283200	-1.30856500
O	2.47277000	0.87176000	-2.40919700
C	0.71500300	6.15381900	-1.22076900
N	1.48956800	5.34964300	-0.27338200
C	1.26470100	4.07232000	0.09217900
N	0.14981400	3.43138700	-0.21996900
N	2.23894300	3.47813600	0.84803600
C	-3.51699500	6.01665400	-0.06160900
C	-3.16964900	4.66767400	-0.73504000
O	-1.95395000	4.59403400	-1.15816500

O	-4.02820600	3.76576100	-0.82749400
Ce	3.62249600	-0.09287700	-0.40538600
N	-3.43375200	-0.57899600	1.07843300
C	-3.87957900	0.24202600	2.07565800
C	-5.34291700	0.36871300	2.37220100
O	-5.70044900	0.88986900	3.42669000
C	-2.79505100	0.89316800	2.65031400
C	-1.65178600	0.46074100	1.94705900
C	-2.06746400	-0.48437600	0.96111400
C	-0.28176500	0.82341000	2.07254900
O	0.04261300	1.71419600	3.04512400
C	0.64676900	0.28330500	1.20473400
O	1.95413800	0.64394900	1.24533000
C	0.22538100	-0.67002700	0.21522600
N	1.19909000	-1.11471400	-0.59202600
C	0.94553300	-2.06743200	-1.49567000
C	2.10595100	-2.43808300	-2.33010000
O	3.22081600	-1.93402800	-2.21757300
C	-0.30670000	-2.67966300	-1.59801300
C	-1.35594900	-2.21838500	-0.80380300
C	-2.65062500	-2.94902200	-0.91389200
O	-3.77048600	-2.47192200	-0.87757200
C	-1.15263800	-1.13588000	0.10096700
H	-11.06632100	-0.13586200	1.98119300
H	-10.85766600	1.24060100	0.90212000
H	-7.66177900	-1.51203000	-4.01215600
H	-7.86880800	-2.66131500	-2.67893200
H	-7.57002900	-0.83095700	-1.34950300
H	-4.47552700	0.83435100	-3.35325000
H	-5.61513800	-0.11610300	-4.21003900
H	-5.03842200	1.58151200	-1.20719800
H	-5.90141100	0.45787200	-0.17098500
H	-0.34840400	-6.36806900	2.04784700
H	-1.43000800	-7.50240400	1.18728000
H	-0.85698500	-5.55493200	-0.08640700
H	0.78999500	7.20204500	-0.92158600
H	-0.33897300	5.85903700	-1.20667500
H	2.35819400	5.74094300	0.05940900
H	0.06846700	2.46358500	0.06916700
H	-0.74797600	3.88476000	-0.61331800
H	3.18666200	3.66704000	0.53765600
H	2.13759500	2.46958400	0.98941400
H	-4.57000900	6.06022500	0.22752200
H	-3.28187500	6.83782700	-0.74675600
H	4.53305100	-4.40762300	2.00649300
H	4.44795900	-3.22769400	3.33473600
H	9.14072200	0.38682600	-1.65415800
H	7.85510500	1.45451400	-2.28345100
H	6.77509400	-0.87799100	0.78789200
H	6.07610300	0.77886700	4.01750300
H	5.35561800	2.34153500	4.41621200
H	1.43199700	3.42187900	-2.83172300
H	2.41995100	2.83508700	-4.17173000
H	-4.04106200	-1.11860100	0.47476500
H	-2.84852500	1.60830500	3.45584800
H	-0.44755900	-3.51346300	-2.27278900
H	-11.32809200	-0.33724500	0.21538600
H	-2.88967600	6.13909800	0.82886200
H	1.10813400	6.04080500	-2.23805300
H	3.09276700	4.04722400	-3.02958100
H	8.06201100	-0.15640000	-2.97012300
H	6.52162500	2.21777300	3.05461400
H	5.98616900	-3.47615600	2.50391800
H	0.32227100	-7.79481400	1.21231100
H	-6.26970300	-2.47443700	-3.43139500
O	-2.46255900	-4.27999700	-1.10593200
H	-3.34072300	-4.69123600	-1.17762200
O	-6.10237700	-0.10529300	1.42279100
H	-7.49674300	0.03506500	1.66613200
O	1.85176700	-3.39227600	-3.23940000

H	2.67749700	-3.56930600	-3.71987700
H	8.45293800	-0.45245000	0.41777500
C	0.93700000	-3.75795000	1.59806300
H	1.45215300	-4.10199400	0.68673100
H	1.59060400	-3.43090500	2.43334300
H	0.99106600	1.91239300	3.01235600
O	-0.27330400	-3.73808400	1.69721600
H	2.80006900	0.74997200	2.52811000

Table S7. INT2_B coordinates

C	10.47060700	0.09000400	1.59447800
C	9.04945500	0.19078700	1.08978400
O	8.20656400	-0.45549000	1.84437500
O	8.75786100	0.82156800	0.06166100
C	7.61449100	2.36654400	-2.94092100
N	7.11267800	1.25604200	-2.11672700
C	6.21458600	0.33894000	-2.50167400
N	5.75705500	0.35036400	-3.78917100
N	5.72944300	-0.55114500	-1.64257900
C	0.36853900	7.43110500	0.43120900
O	0.99821200	6.79639900	-0.67984500
C	-6.68250000	2.51208700	0.45332200
C	-5.65255900	1.57669500	-0.14051700
O	-5.94148200	0.94265100	-1.21436200
O	-4.52710100	1.42599400	0.41371000
C	-6.14272100	-1.38080700	-5.37951300
C	-5.97616600	-0.67461400	-4.05401800
O	-4.89411100	-0.77723700	-3.43957500
N	-6.99866000	0.04564600	-3.57422800
C	-6.20326300	-2.69308800	1.74380100
C	-4.86343700	-2.17903800	1.26360800
O	-4.75549000	-1.91588200	0.00893900
O	-3.91986200	-2.04002300	2.08479500
C	-0.37960700	-2.25586100	-3.48584700
C	-1.59571900	-1.58757300	-2.89152400
O	-2.48752200	-2.29335000	-2.29928900
O	-1.71112400	-0.32366700	-2.93721800
C	-0.58079700	-5.64601700	-2.49738300
N	-1.49329000	-4.85541000	-1.67829400
C	-1.25773600	-4.57124500	-0.38337500
N	-0.08995000	-4.89990700	0.17994000
N	-2.25338800	-4.01279900	0.33546700
C	3.42690200	-5.61371500	-0.70582300
C	3.10598500	-4.18471200	-1.22558400
O	1.85398200	-3.89160800	-1.26021400
O	4.03663900	-3.42466000	-1.57797500
Ce	-3.66930000	-0.25276300	-1.32631500
N	3.15704700	0.07188800	0.81274000
C	3.61008200	-1.02281600	1.48935500
C	5.07424300	-1.24207500	1.72094500
O	5.42382400	-2.13855600	2.48411700
C	2.51235200	-1.77402600	1.89782800
C	1.36983100	-1.09020900	1.45348100
C	1.79170300	0.08405900	0.76515800
C	-0.01829400	-1.45317900	1.58404100
O	-0.43483100	-2.52014300	2.02249600
C	-1.02164100	-0.34196900	1.18188700
O	-2.20585800	-0.87072600	0.69736100
C	-0.48292700	0.64728200	0.13645300
N	-1.40076400	1.20575700	-0.62009900
C	-1.07249100	2.24515800	-1.42252400
C	-2.15966100	2.70741200	-2.30656000
O	-3.25675600	2.16225400	-2.38433300
C	0.19302700	2.80184100	-1.39849400
C	1.21448800	2.18698100	-0.64874400
C	2.54843300	2.84478800	-0.70440200
O	3.64452300	2.33297000	-0.55685400

C	0.91659600	1.00032100	0.08868700
H	10.51534000	0.42053300	2.63637900
H	10.78272200	-0.95907800	1.57454900
H	8.27743300	2.01769000	-3.74378200
H	8.18981200	3.02073600	-2.28570500
H	7.58067000	1.09372700	-1.20872900
H	5.11960900	-0.38319900	-4.06520300
H	6.36079800	0.70118300	-4.51725200
H	5.16702000	-1.35354300	-1.92959700
H	5.86703600	-0.45444200	-0.61572500
H	-0.58878400	6.95641400	0.69152400
H	1.01225200	7.44337800	1.32280800
H	1.19634600	5.87541100	-0.46397500
H	-0.31931300	-6.56410700	-1.96441400
H	0.34370400	-5.10702500	-2.72886800
H	-2.08746400	-4.16644700	-2.14291300
H	-0.00333400	-4.62886500	1.15267800
H	0.77887100	-4.66090200	-0.40725900
H	-3.06762600	-3.67693100	-0.16354900
H	-1.98579100	-3.45826200	1.14320900
H	4.50314400	-5.79667100	-0.69796400
H	2.92836200	-6.35226600	-1.34329500
H	-6.31732000	2.95363300	1.38149300
H	-7.61002300	1.96147500	0.64014500
H	-7.12417300	-1.22786100	-5.83541600
H	-5.97618300	-2.45051700	-5.22381200
H	-6.87615600	0.51780500	-2.67006500
H	-6.97193700	-1.94581000	1.52095300
H	-6.17995800	-2.90088300	2.81406600
H	0.30265700	-2.57576100	-2.68500400
H	0.14808600	-1.56289700	-4.14331000
H	3.76323100	0.76162000	0.37614300
H	2.55572600	-2.69868200	2.45144300
H	0.40332700	3.68468200	-1.98800900
H	11.14132200	0.68927200	0.97747000
H	3.03170400	-5.73378900	0.30899700
H	-1.10113400	-5.91228400	-3.42031800
H	-0.67131700	-3.15012400	-4.04338600
H	-5.36467900	-1.02521500	-6.06115800
H	-6.46334800	-3.60003800	1.18887100
H	-6.91107000	3.30220400	-0.26963900
H	0.17172400	8.46439100	0.13447800
H	6.78428800	2.93449400	-3.37350000
O	2.44376300	4.17559000	-0.96945900
H	3.34569600	4.53314500	-1.03337900
O	5.83275500	-0.39948000	1.07708000
H	7.21144400	-0.43511600	1.50528300
O	-1.84985300	3.79011900	-3.03837400
H	-2.61727200	4.00082500	-3.59603500
H	-7.87148000	0.10783100	-4.07365000
C	-2.28318800	1.25990600	2.51180300
H	-3.22939200	0.71770700	2.47287900
H	-2.16364200	1.74518000	3.48175300
H	-2.28329000	2.00757200	1.71264600
O	-1.16569200	0.35964400	2.40977700
H	-2.78434000	-1.38046000	1.40012000

Table S8. TS3_A coordinates

C	-10.68829200	-1.18870200	-1.96581300
C	-9.37132000	-0.69760600	-1.40761000
O	-8.37653500	-0.88527400	-2.22577100
O	-9.28542900	-0.20789100	-0.26986100
C	-8.45466700	1.67955300	2.58320600
N	-7.71142800	0.69921600	1.78490300
C	-6.63677900	0.02168700	2.20147000
N	-6.21910100	0.15089700	3.49249300
N	-5.95057600	-0.74741100	1.36022200

C	-2.21629900	7.97565900	-0.72776800
O	-2.46227700	6.60009700	-0.44035900
C	4.47457900	3.72570800	-2.58849000
C	4.25794100	2.54830700	-1.66175100
O	5.26105100	2.01028100	-1.08128700
O	3.07988400	2.13118100	-1.44706000
C	8.21948400	-1.17111700	1.05494700
C	7.27810800	-0.15840300	0.45154200
O	6.21186500	0.12723000	1.03610600
N	7.61290400	0.43571700	-0.70257800
C	5.30213600	-2.08348700	-3.82446600
C	4.38929100	-1.59676400	-2.72630200
O	4.84800200	-0.98767600	-1.73159200
O	3.12685800	-1.83521000	-2.88474300
C	3.78970700	-2.31990500	3.78335400
C	3.75330700	-1.48039000	2.52299900
O	3.98053600	-2.02224300	1.38910700
O	3.46976700	-0.24641200	2.62199100
C	1.13202800	-4.61736400	2.54004400
N	2.04215200	-4.34376700	1.44385000
C	1.59997700	-4.23142500	0.15579000
N	0.34158000	-4.40173400	-0.14732800
N	2.61095200	-4.01613400	-0.76676400
C	-2.75366900	-5.38950000	0.64825500
C	-2.81907500	-3.94489900	1.14709000
O	-1.69475500	-3.24236900	1.06518100
O	-3.84576100	-3.44090600	1.59213600
Ce	3.85911100	0.31483400	0.23346000
N	-3.36580200	0.04587800	-0.83196100
C	-3.77697900	-0.92043400	-1.70714200
C	-5.21659600	-1.14239100	-2.05692100
O	-5.48273200	-1.94347200	-2.95223600
C	-2.66675200	-1.60273500	-2.18852200
C	-1.54000000	-1.02959900	-1.56824500
C	-1.99264100	0.02045500	-0.71066000
C	-0.15900700	-1.33465400	-1.69933300
O	0.18136800	-2.33485700	-2.56928800
C	0.74926300	-0.65166500	-0.92353900
O	2.07935400	-0.90422000	-0.94504900
C	0.30852300	0.41140400	-0.06204700
N	1.29175500	1.03393100	0.60360600
C	1.01911000	2.10278300	1.35837900
C	2.20814700	2.75467100	1.94348600
O	3.36410800	2.45421500	1.65140600
C	-0.26762400	2.63714100	1.45457100
C	-1.32530500	2.02269500	0.78158000
C	-2.63021000	2.75955900	0.79771500
O	-3.74741800	2.27857300	0.70526300
C	-1.09515800	0.82467600	0.03291700
H	-10.81584000	-0.83707200	-2.99312100
H	-10.66820400	-2.28351600	-1.99679600
H	-9.01842200	1.20726500	3.39909300
H	-9.16768200	2.16661900	1.91777200
H	-8.15080700	0.38479300	0.90003900
H	-5.39834700	-0.35644600	3.78724700
H	-6.88211100	0.39510400	4.21128500
H	-5.25094300	-1.41086000	1.68212100
H	-6.05712500	-0.64583400	0.32579600
H	-1.59989900	8.45312200	0.04810400
H	-1.72311500	8.10815000	-1.70103500
H	-1.61898500	6.11692700	-0.49080400
H	1.72057100	-4.83056400	3.43572700
H	0.52815100	-5.49663900	2.30192300
H	2.88063400	-3.81213600	1.64824600
H	0.19165100	-4.32761500	-1.15142000
H	-0.90951600	-3.75190700	0.63210200
H	3.42873000	-3.54356700	-0.39715900
H	2.30867400	-3.61108200	-1.64057600
H	-3.65433400	-5.92098600	0.95604600
H	-1.85515700	-5.89146500	1.01630700

H	4.20795300	4.64400100	-2.05236300
H	3.81826600	3.64662000	-3.45908100
H	9.13173200	-1.32072500	0.47194500
H	7.68183400	-2.11971800	1.14450500
H	6.93491600	1.07617700	-1.12275700
H	5.00691300	-1.62018800	-4.77113400
H	5.18459100	-3.16514900	-3.94229000
H	2.76061500	-2.55663000	4.07604800
H	4.23937300	-1.74644500	4.59717700
H	-3.97658900	0.69088700	-0.34394600
H	-2.69447700	-2.41352500	-2.89889100
H	-0.42687200	3.55567000	2.00171300
H	-11.51818900	-0.85782500	-1.33949100
H	-2.69498200	-5.38378500	-0.44563300
H	0.44750300	-3.78330800	2.74677300
H	4.33916000	-3.25041500	3.62404300
H	8.48153600	-0.84452100	2.06511100
H	6.34022100	-1.84085500	-3.59591800
H	5.51856700	3.79405900	-2.89929800
H	-3.18678800	8.47806400	-0.75986300
H	-7.77942700	2.43709700	2.99416600
O	-2.41780000	4.08589800	0.90157400
H	-3.24839700	4.58604700	0.81291600
O	-6.06903600	-0.46621300	-1.33502700
H	-7.42277200	-0.69259300	-1.82514300
O	1.93474000	3.74485400	2.80556600
H	2.78102500	4.11767100	3.10486300
H	8.43985400	0.16652000	-1.21121000
H	1.07116400	-2.20363500	-2.93136000
H	2.54353600	-1.38673500	-1.91427000
C	0.39114700	4.18142700	-1.33000800
H	-0.45264000	3.52784400	-1.62424300
H	1.40641400	3.77717400	-1.49172600
O	0.19524500	5.29582500	-0.87976800

Table S9. TS3_B coordinates

C	10.47065900	0.08998900	1.59454400
C	9.05517200	0.29408700	1.10256700
O	8.16440200	-0.16536500	1.93486900
O	8.81034800	0.83817100	0.01447800
C	7.61445300	2.36658200	-2.94097000
N	7.10257200	1.27931800	-2.09624900
C	6.17948900	0.37935700	-2.45679100
N	5.70347500	0.38118600	-3.73514400
N	5.69576000	-0.49178900	-1.57594600
C	0.36851900	7.43090800	0.43115300
O	1.09291500	6.49406400	-0.36709300
C	-6.24586200	1.30574200	1.70052700
C	-5.47454700	0.57514200	0.62469700
O	-6.07482700	0.25254600	-0.46077200
O	-4.26062100	0.27196800	0.80516200
C	-7.68880700	-1.86417000	-4.45894000
C	-7.05862300	-1.22931100	-3.24143900
O	-5.83539800	-1.37667900	-3.03817800
N	-7.83995100	-0.53945700	-2.40030800
C	-6.64600200	-4.27150300	-1.26947800
C	-5.35100000	-4.12310300	-0.48737300
O	-4.94302600	-2.91489300	-0.28523600
O	-4.73720300	-5.13985000	-0.10663900
C	-1.75303500	-2.95728200	-4.63320200
C	-2.55200700	-2.31366900	-3.52169400
O	-3.07115100	-3.05076900	-2.61236700
O	-2.67415500	-1.05404900	-3.48505300
C	-0.58082300	-5.64578800	-2.49733300
N	-1.50791200	-5.13270000	-1.49567500
C	-1.18542800	-5.01810600	-0.20517600
N	0.05345800	-5.25943800	0.21162800

N	-2.18229300	-4.58473400	0.64746300
C	3.42693400	-5.61376900	-0.70583300
C	3.02999900	-4.18982200	-1.16071800
O	1.76512300	-3.92983300	-1.09127600
O	3.89878800	-3.39009700	-1.55920100
Ce	-4.04102300	-1.05271800	-1.32727600
N	3.07878500	-0.06246100	0.57869800
C	3.60522400	-0.96320900	1.45520200
C	5.05313300	-0.94625900	1.84317300
O	5.41666900	-1.65584900	2.77948100
C	2.59672900	-1.82918000	1.87464700
C	1.42979700	-1.43663300	1.20641500
C	1.73714600	-0.30769400	0.39266200
C	0.11133200	-1.99626300	1.23317000
O	-0.28077000	-2.90021900	2.07113200
C	-0.87142800	-1.37912100	0.41055900
O	-2.06758300	-1.93902200	0.20517200
C	-0.56208600	-0.20436100	-0.39273600
N	-1.59187200	0.29146900	-1.06749300
C	-1.44618100	1.43687800	-1.74822100
C	-2.66302600	1.91412400	-2.43502800
O	-3.76393900	1.38122200	-2.33760800
C	-0.24401700	2.14618500	-1.74218700
C	0.86453900	1.64517100	-1.05895300
C	2.09287600	2.49176000	-1.08734800
O	3.24905300	2.10777500	-1.05510400
C	0.76103800	0.39668600	-0.36156400
H	10.58408400	0.53870100	2.58596200
H	10.65942400	-0.98286800	1.70198000
H	8.24831400	1.99281600	-3.75633600
H	8.22322300	3.01032300	-2.30587700
H	7.59598000	1.10879000	-1.20249800
H	5.02516000	-0.32170900	-3.99043400
H	6.28955500	0.72563600	-4.47979600
H	5.10972300	-1.27909100	-1.84941000
H	5.82408300	-0.35381300	-0.55066700
H	-0.67199800	7.54046300	0.09415300
H	0.37353700	7.15681900	1.49533000
H	0.68844900	5.62000800	-0.29221500
H	-0.11044400	-6.55992800	-2.12491900
H	0.20895400	-4.92436600	-2.73164700
H	-2.31141800	-4.59720700	-1.83690900
H	0.18984800	-5.11680300	1.20557400
H	0.88623800	-4.80126200	-0.40679100
H	-3.17318400	-4.84147400	0.40266000
H	-1.95237600	-4.66995600	1.63025500
H	4.49509800	-5.78480300	-0.85106900
H	2.84959600	-6.35773100	-1.26524700
H	-5.56531000	1.72514200	2.44309800
H	-6.91809000	0.59441700	2.19341800
H	-8.72492500	-1.55724200	-4.62213100
H	-7.65341800	-2.95149000	-4.33673600
H	-7.41250700	-0.12480800	-1.56231300
H	-7.38891600	-3.55130200	-0.91515100
H	-7.02920400	-5.29051000	-1.20031000
H	-0.72090800	-3.08468600	-4.28633800
H	-1.74306300	-2.31609400	-5.51656500
H	3.61546400	0.65003600	0.09717900
H	2.72673000	-2.64945200	2.56290100
H	-0.17837000	3.09036000	-2.26621300
H	11.18581000	0.52733400	0.89692500
H	3.18494700	-5.73983300	0.35568000
H	-1.15375700	-5.87953400	-3.39592700
H	-2.15541700	-3.94357400	-4.87574600
H	-7.08705500	-1.61229700	-5.33554500
H	-6.42565500	-4.03537900	-2.31671900
H	-6.86204200	2.09123200	1.25439400
H	0.87197700	8.39405900	0.31902900
H	6.79035400	2.95477300	-3.35780100
O	1.78850400	3.80866300	-1.17944700

H	2.61461400	4.32187100	-1.20964200
O	5.79099800	-0.14522200	1.12590800
H	7.17295100	-0.12598400	1.58937500
O	-2.47841200	3.02245800	-3.17383500
H	-3.33297700	3.25139900	-3.57599500
H	-8.82594900	-0.43207500	-2.57766700
C	-1.65784100	-1.06251100	3.16212400
H	-1.08508100	-2.20722000	2.87041700
H	-1.17555400	-0.94641500	4.14858900
H	-2.73586200	-1.29151900	3.17336000
O	-1.22546400	-0.27602900	2.22045000
H	-2.11690100	-2.91019600	0.49787600

Table S10. EP coordinates

C	-10.65572100	-1.44361100	-1.67583100
C	-9.37276300	-0.79012500	-1.21376900
O	-8.44267300	-0.79771000	-2.12654000
O	-9.25092800	-0.32774500	-0.06921800
C	-8.29953300	1.42246700	2.81233600
N	-7.56883900	0.51788900	1.91780000
C	-6.47339700	-0.18293200	2.23512300
N	-6.00208700	-0.15814600	3.51426200
N	-5.81829100	-0.87065200	1.30355800
C	-2.32259600	7.87579100	-0.67407500
O	-2.38983300	6.57691400	-0.08634100
C	3.87708800	3.47423900	-2.80279700
C	3.78702500	2.47158000	-1.67387700
O	4.82270900	2.27236400	-0.94509600
O	2.71775600	1.83190900	-1.46605900
C	8.65183000	-0.22504600	0.55162700
C	7.38637800	0.49827300	0.15722700
O	6.30117300	0.16280400	0.67956600
N	7.45933000	1.48396400	-0.74521600
C	4.70566100	-1.18878800	-4.23998800
C	3.98654500	-1.26398700	-2.91120600
O	4.52008000	-0.74180200	-1.89107100
O	2.84828800	-1.84927200	-2.88576700
C	4.41223400	-2.60540600	3.62672200
C	4.17571200	-1.70499300	2.43176900
O	4.24705800	-2.17914500	1.25003900
O	3.88256400	-0.48586000	2.64004000
C	1.39952100	-4.68062300	2.32083600
N	2.31001100	-4.35080600	1.23376600
C	1.93396900	-4.25194400	-0.05511100
N	0.67169500	-4.39380400	-0.42251200
N	2.94232500	-4.08819400	-0.96336000
C	-2.54233400	-5.51081500	0.57597200
C	-2.46178600	-4.01831600	0.98675700
O	-1.28649100	-3.48760600	0.91657800
O	-3.49062900	-3.42327900	1.36555400
Ce	3.84641800	0.26533600	0.28909400
N	-3.41284900	0.25520900	-0.77729100
C	-3.83536600	-0.58527800	-1.76808100
C	-5.27194600	-0.69712300	-2.18469000
O	-5.52306900	-1.22609300	-3.26700500
C	-2.74381300	-1.28196900	-2.27272300
C	-1.62280300	-0.85850000	-1.53723400
C	-2.05099600	0.13405200	-0.60212300
C	-0.26293300	-1.26115300	-1.61528200
O	0.07220900	-2.18390300	-2.56118900
C	0.62382600	-0.72511400	-0.71330600
O	1.93582800	-1.12141200	-0.63901200
C	0.23743700	0.32574200	0.17651200
N	1.24178300	0.87627900	0.87519900
C	1.00075600	1.93188700	1.66307800
C	2.21079900	2.53829600	2.25056800
O	3.35519600	2.25643300	1.89283500

C	-0.26402300	2.51267600	1.77312500
C	-1.33559200	1.98565000	1.04935500
C	-2.60184800	2.78874700	1.06518400
O	-3.73420400	2.36875500	0.89261600
C	-1.14311900	0.82173000	0.23955400
H	-10.92013500	-1.09053700	-2.67613600
H	-10.48987300	-2.52419500	-1.74482300
H	-8.84719800	0.87913500	3.59445600
H	-9.02466900	1.96414000	2.20467600
H	-8.04329900	0.26004800	1.03466100
H	-5.18156300	-0.70612500	3.72889400
H	-6.64394600	0.00444400	4.27460100
H	-5.08916200	-1.54470200	1.53308600
H	-5.98213200	-0.66970100	0.29324400
H	-1.32084300	8.31805000	-0.57491500
H	-2.60156500	7.86045500	-1.73704500
H	-1.77552400	5.98100400	-0.54354900
H	1.99451100	-4.85012500	3.22037200
H	0.85730400	-5.59944600	2.08043500
H	3.16685000	-3.85117900	1.45408800
H	0.48995000	-4.28944600	-1.41382700
H	-0.20129300	-4.09004000	0.22739200
H	3.82736400	-3.78403600	-0.57442700
H	2.72013400	-3.61102300	-1.82943500
H	-3.56853100	-5.87640300	0.64321700
H	-1.89489400	-6.10572500	1.23014400
H	2.89386300	3.66665700	-3.23489400
H	4.53672000	3.07201600	-3.57998400
H	9.54840600	0.16920200	0.06689300
H	8.53597700	-1.28270300	0.29722000
H	6.58960900	1.96265500	-1.01257200
H	4.23322400	-0.40260500	-4.84025500
H	4.59410600	-2.12850100	-4.78631500
H	3.45258400	-2.79434200	4.12093900
H	5.05374800	-2.09514100	4.34963400
H	-4.00694000	0.88065000	-0.24422300
H	-2.78357900	-2.01822400	-3.05938800
H	-0.39267200	3.40991300	2.36225700
H	-11.46269200	-1.24396200	-0.96965800
H	-2.17553400	-5.63646700	-0.44880300
H	0.65999000	-3.89248500	2.50077800
H	4.86226800	-3.55451700	3.32785400
H	8.76548000	-0.15977700	1.63722700
H	5.75929300	-0.94338200	-4.09849500
H	4.32376200	4.40480000	-2.44123100
H	-3.03513000	8.50738500	-0.13684100
H	-7.61712200	2.14240600	3.27564900
O	-2.33380700	4.08823100	1.28052300
H	-3.12553500	4.64019600	1.14870400
O	-6.12859200	-0.21340100	-1.33000200
H	-7.49669600	-0.49433300	-1.78965900
O	1.97833800	3.46858600	3.18694000
H	2.83770900	3.81927000	3.47587600
H	8.33837300	1.74317300	-1.16405000
H	0.98441400	-2.05006400	-2.88148700
H	2.27567700	-1.47421700	-1.53335100
O	-0.79096300	4.44903500	-1.57336900
C	-0.16507100	3.40738300	-1.62380400
H	-0.47497600	2.57970600	-2.28973500
H	0.74148200	3.22726100	-1.02139500

Paper II

"Reaction Mechanism of Low-Spin Iron (III)-and Cobalt (III)-Containing Nitrile Hydratases: A Quantum Mechanics Investigation"

Mario Prejanò, Tiziana Marino, Jose Carlos Madrid Madrid,
Nino Russo, Marirosa Toscano
Inorganic Chemistry **2017**, *56(21)***,** 13390-13400.

doi: 10.1002/chem.201700381

Inorganic Chemistry

Article

Cite This: Inorg. Chem. 2017, 56, 13390-13400

pubs.acs.org/IC

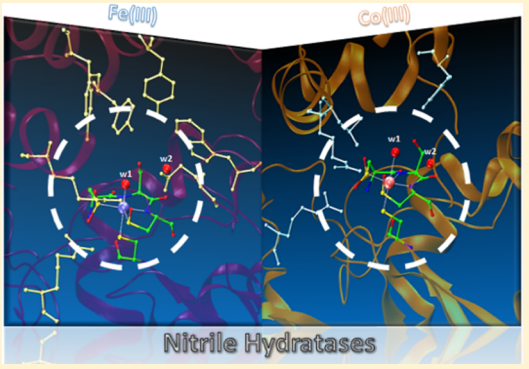
Reaction Mechanism of Low-Spin Iron(III)- and Cobalt(III)-Containing Nitrile Hydratases: A Quantum Mechanics Investigation

Mario Prejanò, Tiziana Marino,* Carmen Rizzuto, Josè Carlos Madrid Madrid, Nino Russo, and Marirosa Toscano*

Department of Chemistry and Chemical Technologies, Università della Calabria, Via P. Bucci, I-87036 Arcavacata di Rende, Italy

Supporting Information

ABSTRACT: To elucidate the catalytic mechanism of cobalt(III)−benzonitrile and iron(III)−-pivalonitrile hydratases, we have performed at density functional level a study using the cluster model approach. Computations were made in a protein framework. Following the suggestions given in a recent work on the analogous enzyme Fe(III)−NHase, we have explored the feasibility of a new working mechanism of examined enzymes. According to our results, after the formation of enzyme substrate complex, the reaction evolves toward product in only three steps. The first one is the nucleophilic attack, led by the −OH group of the αCys113−S−OH on the nitrile carbon atom, followed by the amide formation and by the enzyme restoring phase that our computations indicate as the most expensive step from the energetic point of view in both catalytic processes.



Downloaded via UNIV DELLA CALABRIA on November 24, 2018 at 19:28:20 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Nitrile-containing compounds are fundamental components of the metabolic cycle of the upper plants and of many soil microorganisms[1,2] as they act as growth hormones during germination and are essential constituents of agents for the chemical protection of plants from herbivores.[3,4] Organic nitriles are widely used at the industrial level for the production of plastics, fibers, pesticides, water treatment reagents, and other chemical products.[5] Several anthropogenic nitriles are, however, toxic and have carcinogenic and teratogenic properties.[6−11] Industries generate huge amounts of chemical waste containing nitriles, which are often difficult to dispose.[12] Modern biotechnological methods based on the defined "enzymatic" approach study the specific chemical reactions catalyzed by the enzyme synthesized within a microorganism. The interest in the study of the catalytic activity of the enzymes is aimed to clarify the molecular mechanisms and their reactivity addressing the attention mainly to the active site and its electronic properties.[13] In a microbial transformation, enzymes act as biocatalysts and in addition to natural substrates, many of these enzymes can use structurally related substrates and be able to catalyze the reaction equally. In this sense, microbial transformation processes represent a specific category of chemical synthesis.[14] In the chemical synthesis, one example of proteins studied on a large scale is the enzyme nitrile hydratase (NHase), purified and characterized over the years by different microorganisms such as, *Pseudomonas putida*,[15] *Rhodococcus rhodochrous* J1[8,16] and *Pseudonocardia Thermophila* JM 3095.[17] Microbial enzymes that catalyze the hydrolysis of organic nitriles are divided into two main classes:[5] nitrilases and nitrile hydratases. Nitrilase enzymes hydrolyze organic nitriles in the corresponding carboxylic acids and ammonia and do not use metal cofactors for catalysis, instead, Nitrile hydratases hydrolyze organic nitriles in their amides[18] which eventually, in a second step, can be transformed by the amylases in carboxylic acids and ammonia.[19]

In 1980, Asano et al.[20] studying the microbial degradation of acetonitrile, discovered the existence of an enzyme able of catalyzing the hydration reaction of nitrile in amides at room temperature and physiological pH conditions. The enzyme called nitrile hydratase, or more simply NHase, became part of the enzymatic class EC 4.2.1.84. The enzymes of the nitrile hydratase family are of paramount importance for the production of acrylamide and nicotinamide.[18] In 1985, the industrial production of acrylamide with nonheme iron-based NHase from *Rhodococcus* SP N-774,[21] subsequently replaced by NHase of second generation obtained from *Pseudomonas chlororaphis* B23, was launched in Japan. Currently, the NHase from *Rhodococcus rhodochrous* J1 (Co^{3+} low spin) is used in large scale for the production of acrylamide (30,000 tonnes per year).[22] These enzymes, according to their selectivity toward aromatic and aliphatic substrates,[23] can also be used for the synthesis of other chemical products that are useful at the pharmaceutical level, starting from the NHase contained in the *Pseudomonas putida*.[24] NHases work at low spin multiplicities: ($S = 1/2$) for Fe(III) (Fe-type NHases) and ($S = 0$) for Co(III) (Co-type NHases) ion in a nonheme and a noncorrinoid group, respectively.[25] In biochemistry, cobalt ion is in general found in a corrin ring, such as in vitamin B12. So,

Received: August 18, 2017
Published: October 23, 2017

© 2017 American Chemical Society

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

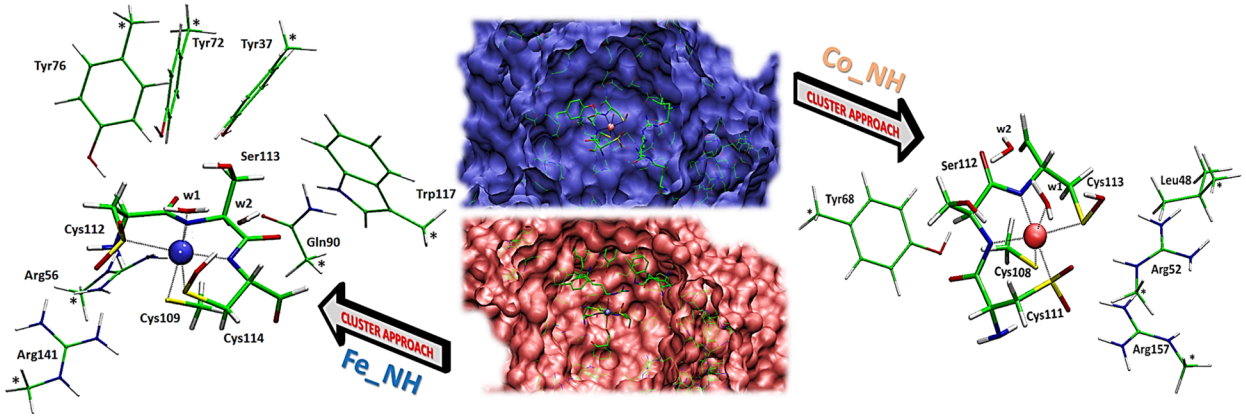


Figure 1. Active site model for Co (III)−NHase (right side) and for Fe(III)−NHase (left side). Stars indicate the atoms kept fixed during the geometry optimizations of all the species intercepted along the potential energy surface.

nitrile hydratase is one of the rare enzyme that uses this metal center in a noncorrinoid manner. Despite the different preferences for the metal, the Fe-type and Co-type NHases exhibit great similarity in the primary and secondary structures and generally they use the same reaction mechanism. The NHase enzymes are tetramer of 92 kDa,[5] and all protein ligands of the metal ion are provided by the α subunit. The metal ion is located in the central cavity at the interface between two subunits α and β and is coordinated in a characteristic "claw setting" to an axial cysteine thiolate, two equatorial peptide nitrogens, an equatorial sulfur atoms belonging to two active site cysteine residues post-translationally modified to cysteine−sulfenic (Cys−SOH) and cysteine−sulfinic (Cys−SO_2H) acids, and a labile axial water molecule.[26] The formation of cysteines post-translationally modified is supported by Mass Spectrometry,[17−20] sulfur K-edge Xray absorption spectroscopy and IR spectroscopy in Fourier transform-FTIR studies.[29,28] The same experiments have indicated that both Cys−SO_2H and Cys−SOH are deprotonated.[29] Asymmetric oxidation of cysteines appears to be essential for the catalytic activity of NHases, because the reconstitution of the enzyme under anaerobic conditions (resulting in a lack of cysteines oxidation) or oxidation of both cysteines to sulfinic acid (Cys−SO_2H), abolishes the catalytic activity of NHase. In addition, the oxygenation of cysteines[25] is crucial in that it removes electron density from the metallic ion enhancing its character of Lewis acid. It was suggested that sulfenate may be protonated and that it is part of a network of strong hydrogen bonds.[30] Enzymes in which Fe(III) and Co(III) metal ions have octahedral geometries tend to be less reactive than those involving ions having pyramidal structures especially if these are at low spin. However, these last are difficult to isolate.[25] Contrary to the Fe(III)−NHases which prefer aliphatic nitriles as substrates, the Co(III)−NHases show greater affinity for aromatic nitriles and are more stable. EPR data are available[30] only for Fe-type NHase, since the presence of a cobalt with d^6 low spin electronic configuration, makes the Co(III)−NHase enzymes not inspected by EPR spectroscopy. Moreover, few other spectroscopic data are available for Co(III)−NHases.[25] It was demonstrated that NHases are good biocatalysts in preparative organic chemistry in that they are capable of hydrating nitriles under physiological reaction conditions,[31,32] so a detailed knowledge of their catalytic mechanism can be particularly useful. With the purpose to understand the role of the metal ions (Fe and Co) inside the NHases and to elucidate the work mechanism of both Fe(III)− and Co(III)−NHases, we have undertaken a comparative study by applying the cluster approach.

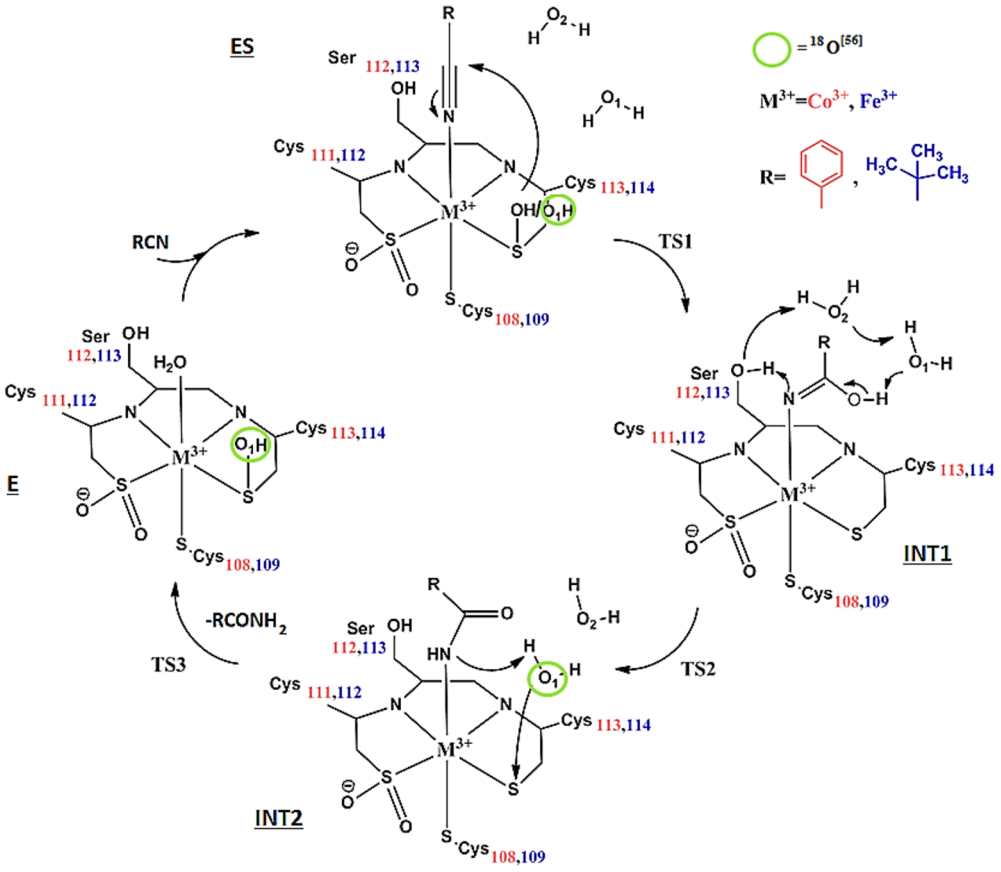
COMPUTATIONAL METHODOLOGY

Active Site Model. The study began with the assembly of a model for the active site of the enzymes following a well-consolidated procedure in the framework of density functional theory known as "cluster approach" and applied in many works devoted to enzyme chemistry.[33−36] This procedure takes into account the fact that normally the enzymatic reaction occurs in a well-defined area of the enzyme (active site). Only the functional groups primarily belonging to lateral chains of first (and sometime of second) shell amino acid residues participate to the reactions. The rest of the enzyme is considered as a matrix, in which the active site is incorporated, which provides to its structural stabilization and solvation. The active site model for our enzymes was obtained starting from the X-ray structure derived from the microorganism *Pseudonocardia Thermophila* JCM 3095 in the case of Co(III)−NHase (pdb code: 1IRE, at a resolution of 1.8 Å)[37] and from *Rhodococcus erythropolis* N771 for the Fe(III)−NHase (pdb code: 2ZPE, at a resolution of 1.48 Å)[38] (see Figure 1). The structures of Fe- and Co-type NHases are very similar although Fe-type NHases only bind Fe(III) and Co-type ones only Co(III). This specificity is regulated by the respective activator proteins[40] which produces the same primary ligands with some small difference concerning the nature of those belonging to the outer shells that in turns is strictly related to the different nature of substrates that can be processed.[39,40]

Co-Type NHase. It contains 116 atoms, has total charge +1, and involves the trivalent cobalt ion, four amino acid residues for each subunit α and β: αCys108, αCys111, αSer112, αCys113 belonging to the first metal ion coordination sphere, and βArg52, βArg157, β Leu48, β Tyr68 constituting its second shell.[37] All these residues were proposed to be important for the reaction mechanism.[37,41] Moreover, two crystallographic water molecules (**w1** and **w2**) are included that in the investigated reaction mechanism, will play an active role in proton transfers steps (TS2 and TS3). At this purpose, in Figure S1 is reported the active site for Fe- and Co-type NHases derived from crystallographic structure where the deeper water molecules are evidenced. All the amino acids are truncated at the α-carbons and hydrogen atoms are added manually. To model the steric effects, the method to lock key coordinates at the periphery of the model where the truncation is made, allows to prevent large artificial movements of the active site groups during the geometry optimizations. As demonstrated earlier,[35,42] this strategy generates few imaginary frequencies whose values are very small and do not influence the zero point energy. Besides, the negligible differences in the description of energetics arising from the use of this approach that makes the system slightly rigid, do not alter the conclusions about the mechanism

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

Scheme 1. Proposed Mechanism for the Nitrile Hydrolysis Catalyzed by Co(III)− and Fe(III)−NHase Enzymes



followed by enzyme.[37,42,43] As far as αCys111, αSer112, and αCys113 amino acid residues are concerned, part of their main chain was left unchanged. Instead, βTyr68, αCys108, the two βArg52 and βArg157 and βLeu48 residues were simulated by a phenol ring, a CH_3S^- ion, a $[CH_3NHC(NH_2)]^+$ guanidinium, and a $CH(CH_3)_3$ group, respectively. The substrate (benzonitrile) was chosen on the basis of the Co(III)−NHase specificity.[19] Its coordination and orientation is similar to that present in the available enzyme−inhibitor complex Co(III)−NHase−PBA, where PBA stands for phenylboronic acid.[41]

Fe-Type NHase. In Figure 1 (left part) is illustrated the active site model employed for the investigation of the Fe-type NHase working mechanism obtained applying the above-described procedure. The model includes Tyr37, Tyr72, Tyr76, Arg56, and Arg141 residues belonging to the β unit and Trp117 and Gln90 belonging to the α one. In particular, the Arg56 is conserved in all known NHases and it is always involved in hydrogen bond network with both oxidized cysteines.[37,44,45] The residues of the inner coordination shell of iron ion are the same of those around cobalt (see Figure 1). As in the case of Co-type enzyme the two water molecules present in the active site of the enzyme are retained in the model (see Figure S1). Thus, the structure around two metal centers appears to be almost the same along with the network of hydrogen bonds in which the arginine residues are implicated with the modified cysteines.[37] The substrate is the pivalonitrile. By applying the truncation as above-described, a cluster of 164 atoms with total charge equal to +1 is obtained.

Technical Details. The Gaussian 09 program package[46] was used for the calculations. Geometry optimizations of all the examined species were carried out by using the M06L[47] exchange-correlation functional which takes into account long-range interactions and the effects of dispersion whose role is important in determining the energetic of an enzymatic process. The 6-31+G(d,p) all-electron basis set was used for all atoms except for the cobalt and iron ions which were described by the SDD pseudopotential and its related basis set.[48] Vibrational frequencies were computed at the same level of theory as the optimizations to have zero-point corrections to the energies (ZPE) and to confirm the nature of the stationary points lying on the potential-energy surfaces. NBO analysis was performed to calculate the Wiberg bond orders which can be useful to clarify the dissociative or associative nature of the hydrolysis step.[49] The effects of the protein environment, were estimated using the Self Consistent Reaction Field (SCRF) SMD approach of Truhlar's and co-workers which is recommended as the preferred model to calculate ΔG of solvation.[50] The dielectric constant value $\varepsilon = 4$, commonly used to simulate the natural surroundings for a protein in quantum-chemical modeling of enzymatic reactions,[51−53] was employed to perform single-point calculations with the larger basis set 6-311+G(2d,2p) on the optimized gas-phase structures.

RESULTS AND DISCUSSION

The biocatalytic process of the nitrile hydrolysis by both Fe(III)− and Co(III)−NHases was extensively studied at the experimental level,[27,29,30,41,54] but the exact catalytic path remains poorly understood. At theoretical level, exhaustive studies were mainly performed on the Fe(III)−NHase enzyme.[29,55−59] On the contrary, as far as Co (III)−NHase is concerned, available theoretical works are addressed to the study of the nucleophile attack on the deprotonate residue αCys113−SO^- or on the water molecule activated by the βTyr68 base on the acetonitrile substrate.[26] Several plausible mechanisms of the catalytic reaction of NHase have so far proposed[25,27,29,55,56] but recent experimental studies support direct coordination of the nitrile substrate to the metal center during the catalytic hydration reaction,[15,54] so following these indications, we have explored the mechanism depicted in Scheme 1 for both types of NHase. The M06L optimized

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

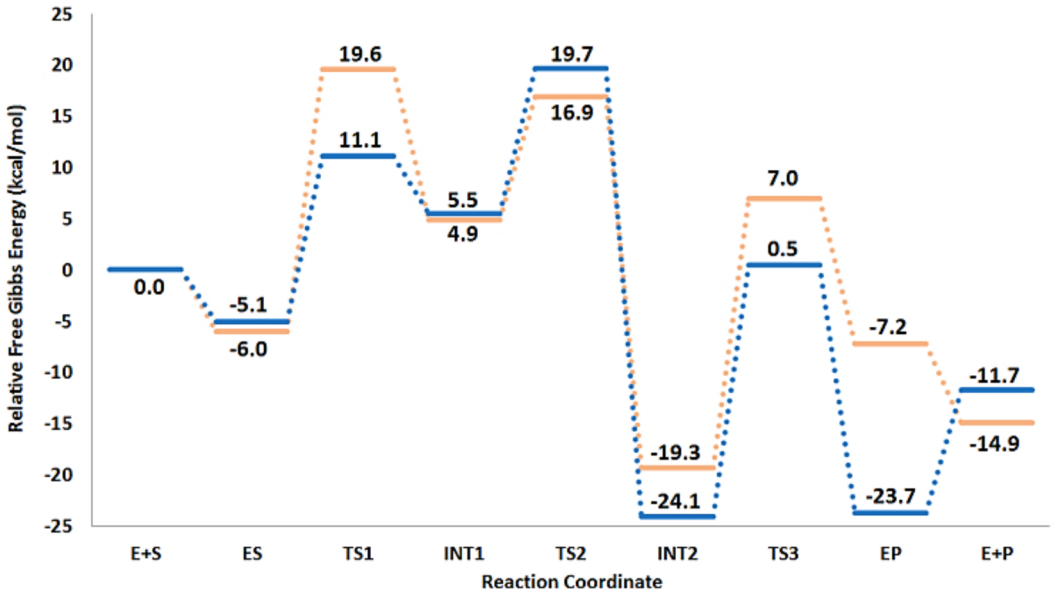


Figure 2. Potential free energy profile for the conversion of benzonitrile to benzamide catalyzed by Co(III)−NHase enzyme (orange line) and by Fe(III)−NHase enzyme (blue line) obtained at M06*L*/6-311+G(2d, 2p) level and in protein environment (ε = 4). Relative free energy values are in kcal/mol.

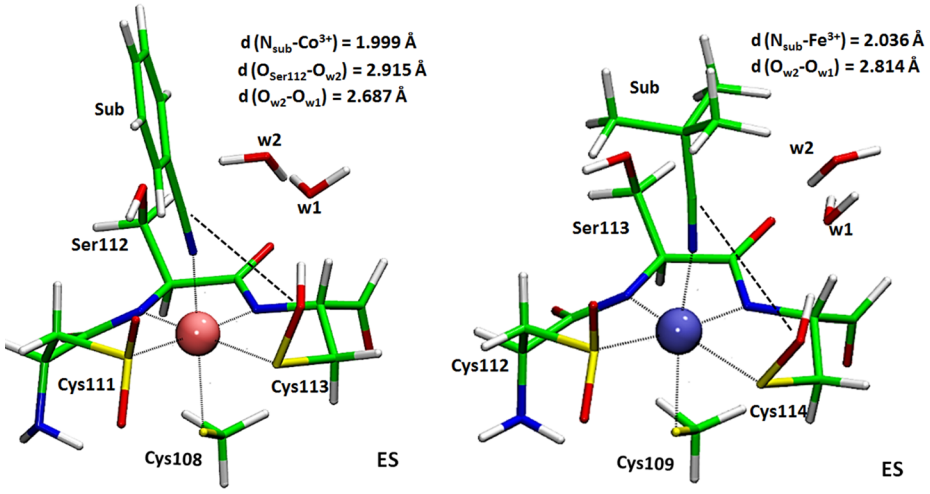


Figure 3. M06*L*/6-31+G(d,p) (SDD for metal ions) optimized structures of Michaelis−Menten complex (ES) for the Co−NHase and Fe−NHase. For clarity, only the amino acid residues of the inner coordination shell of the metal center are retained.

structures of intermediates that are not reported in the main text will be given in Figure S2.

Co−NHase. This work represents the first theoretical investigation of the whole catalytic process for the transformation of benzonitrile into benzamide according to the reaction mechanism resulting by the experimental observations[15] (Scheme 1) and suggested also by a recent work on the analogous enzyme Fe(III)−NHase.[54] As a preliminary step in our study, the lowest-energy spin state of the Michaelis−Menten complex (ES) was determined. Among the three possible values 1, 3, and 5 of 2S+1 multiplicity, computations indicated that the singlet, is the most stable one in agreement with the experimental determination which proposes a spin state for cobalt(III) ion equal to zero.[29,39] The potential energy surface (PES) obtained for the conversion of benzonitrile to benzamide catalyzed by Co(III)−NHase enzyme and obtained in a protein environment is reported in Figure 2. Data in this figure are referred to the sum of the separated reactants energy. When the substrate (S) enters in the catalytic cavity, it causes the displacement of the water molecule present in the apical position of the octahedron generated by ligands around the metal ion giving rise to the ES.

In this complex, the catalytic pocket, created by the amino acid residues belonging to an external coordination sphere of metal ion (βLeu48 and βTyr68), allows the insertion of the benzonitrile substrate that replaces the water molecule and coordinates to the metal ion in apical position with a distance of 1.990 Å (Figure 3). The water molecule (**w1**) reaches a more external coordination sphere at 4.234 Å from Co(III) where it establishes a hydrogen bond with the water molecule (**w2**) present in the active site (1.743 Å). The two water molecules present in the active site are essential since they stabilize the ES complex by guaranteeing a network of hydrogen interactions with the OH group of αSer112 and βTyr68 and with the nitrogen atom of the substrate (distances are 1.955, 1.842, and 2.436 Å, respectively). As can be seen from the PES (Figure 2),

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

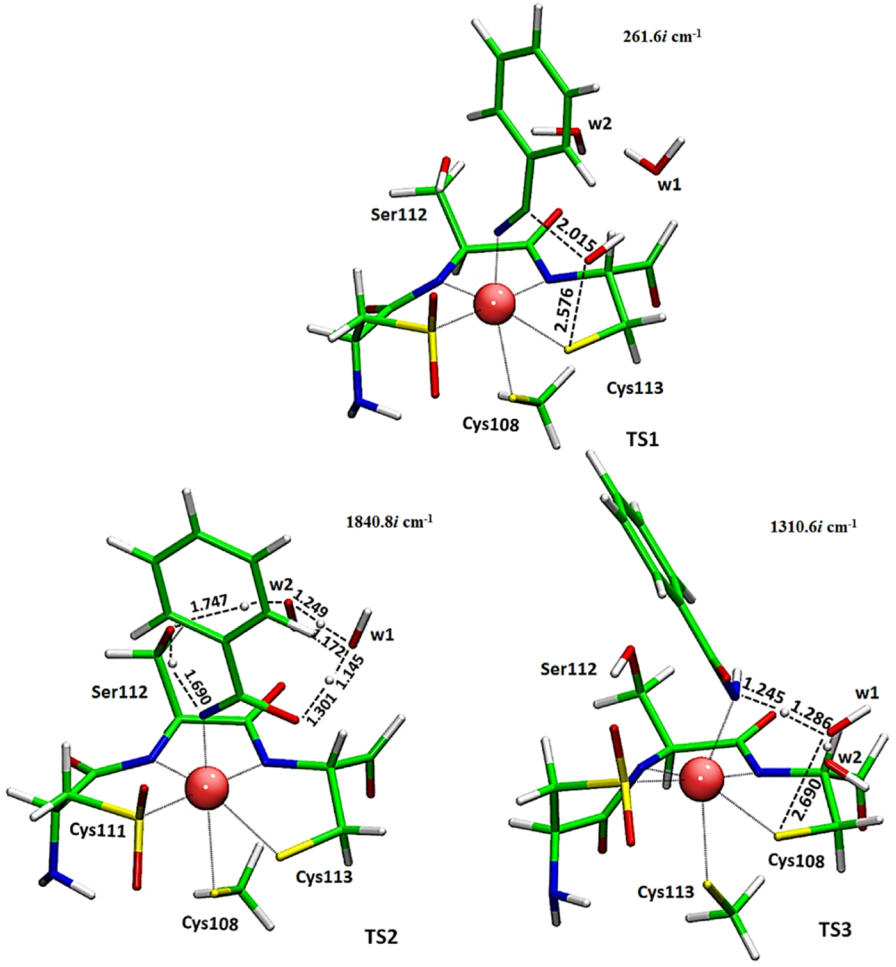


Figure 4. M06*L*/6-31+G(d,p) (SDD for metal ion) optimized structures of the transition states for Co-NH type. For clarity, only the amino acid residues of the inner coordination shell of the metal centers are retained. For each TS imaginary frequencies values are reported.

ES lies at 6.0 kcal/mol below the separate reagents asymptote (E + S). The slight stabilization of the ES complex should allow an inexpensive transition to the next intermediate. In fact, the enzyme should not bind too firmly to the substrate since a too deep energy minimum for the formation of the ES complex could result in a decrease in the efficiency of the enzyme itself.

The novelty of the mechanism followed in this work consists in the fact that the agent that performs the nucleophile attack on the nitrile carbon is no more the −OH group of the serine residue as previously suggested,[56] but the − OH group of the αCys113−S−OH which, as supposed in the most recent work on the enzyme Fe(III)−NHase,[57] is sufficiently polarized to act as nucleophile without the aid of a base that activates it. In fact, the distance between the above-mentioned group and the carbon of the benzonitrile in the complex ES (3.608 Å) is quite suitable because the nucleophile attack can take place. The TS1 transition state describes this process. In its optimized structure, shown in Figure 4, it is possible to see how the linear geometry of nitrile group in the ES complex, has already undergone a deformation of the bonding angle (142.2°). The nucleophile group is now located at 2.015 Å from carbon, whereas its bond with the sulfur atom has lengthened by assuming the value of 2.575 Å (it was 1.712 Å in the ES complex). Moreover, the negative charge generated on the nitrile nitrogen atom after the nucleophile attack makes its coordination bond to the metal ion shorter (1.890 Å) than that present in the ES (1.990 Å). All these changes, both for the substrate and the nucleophile, contribute to the formation of a pseudocyclic five-termed structure in this region according with to what was observed in the crystallographic studies on the Co(III)−NHase-inhibitor (PBA) complex, which is analogous to the present transition state.[18] The value of the imaginary frequency (261.6*i* cm^{-1}) obtained from vibrational analysis correlates well with the simultaneous stretching of the S−OH and OH−N bonds. The TS1 lies at 19.6 kcal/mol above the separate reagents energy. The barrier that must be overcome to continue with the catalytic event is 25.6 kcal/mol.

The lack of previous computational study performed with M06L or similar functionals and mainly the different nature of involved nucleophile, makes a punctual comparison with literature sources impossible. Some previous theoretical works on Fe(III)−NHase have proposed barriers that range from 20.2 to 22.7 kcal/mol depending on the explored mechanism. At the experimental level and always for the iron-containing enzyme, the determined rates indicate even lower barriers.[60]

In INT1 intermediate, which lies at 4.9 kcal/mol with respect to reactants (E + S) and is obtained after the nucleophile attack, the iminol moiety (see Figure S1) conserves the same distance from metal ion (1.890 Å) as in TS1, whereas the oxygen atom of the nucleophile group moves away from the sulfur of the αCys113 at 3.169 Å and forms a bond with carbon whose length is 1.411 Å. It is important to highlight that the −OH of

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

the αSer112 residue appears to be closer to the nitrogen atom of the substrate (1.984 Å) than in the ES complex (2.436 Å), as if what happened during the nucleophile attack, i.e., the hydrogen interaction of serine with the substrate, is served already to prepare the amino acid residue to its next role in the catalytic process. The reaction continues with the formation of the amide through the TS2 transition state shown in Figure 4. The TS2 describes a multiple proton transfer between the αCys113−OH and the nitrogen atom of benzonitrile mediated by the two water molecules that assist the catalytic event. In particular, the water molecule adjacent to the −OH group of the iminol (**w1**) receives a proton (1.133 Å) and gives the remaining one to the nearby water molecule (**w2**) (1.229 Å) which in turn transfers a proton to the − OH group of the αSer112. Then, the αSer112 delivers it to the nitrogen atom of the substrate. Such a concerted transfer of protons is reflected in the value of the imaginary frequency that is 1840.8*i* cm^{-1}. This behavior explored in the present work is in agreement with the experimental findings.[17] The next point on the PES is the intermediate INT2 in which we can observe the result of the transfers described in the TS2 (Figure S1). The benzamide is already formed so that the process of hydration of the nitrile could be considered practically finished. INT2 lies at 19.3 kcal/mol below the separate reactants resulting in an important thermodynamic stabilization of the complex. The benzamide is still coordinated to metal ion with a distance of 1.898 Å. The release of the product will provide the catalysts in its active form. In fact, in the enzymatic catalysis it is essential to restore the enzyme in the active form so that the catalytic cycle can be reactivated. For this reason, we continued the theoretical investigation in order to describe the mechanism of enzymatic restore cycle. To this end, two possible ways were individuated. In the first one, the imidate is transformed in the corresponding amide upon nucleophile attack by a water molecule on sulfur atom of the αCys113. In the second path, the same water molecule leads its attack to the carbon atom of imidate so that the oxygen atom in the product is just derived from the water molecule of the active site. All attempts to simulate this second possibility failed. Instead, for the first mechanism a transition state (TS3) was located which is quite adequate to describe what expected. In fact, the water molecule that in the INT2 was at 4.484 Å from the sulfur of the αCys113, is now at 2.551 Å from it and acts as acid toward the −NH group. Its proton bridges the oxygen atom and the nitrogen atom (Ow—H and H---NH distances are 1.319 and 1.217 Å, respectively) and establishes a further hydrogen bond with the nearby water molecule (1.578 Å). The value of the imaginary frequency of 1310.6*i* cm^{-1} refers clearly to the stretching of the O−H bond, which contributes more than the simultaneous S−O bond one. The barrier to overcome TS3 and to reach the product is of 26.3 kcal/mol (Figure 2). Through the TS3, the reaction evolves toward the EP complex whose optimized geometry is illustrated in Figure 5. As can be noted the benzamide is now only weakly linked to the metal center (Co^{3+}---NH_2 is 3.410 Å) and ready to move away permanently. The water molecule establishing a hydrogen bond with both the βArg52 (2.426 Å) and the product (1.809 Å), contributes to the stabilization of the EP complex lying at 7.2 kcal/mol below the E + S asymptote. The structural characteristics of the EP complex suggest that the restoring of the enzyme will take place rather easily. In Figure S2, we report the energy graph as a function of the shortening of the water−metal bond. From this graph, it is possible to see how the energy tends to become increasingly

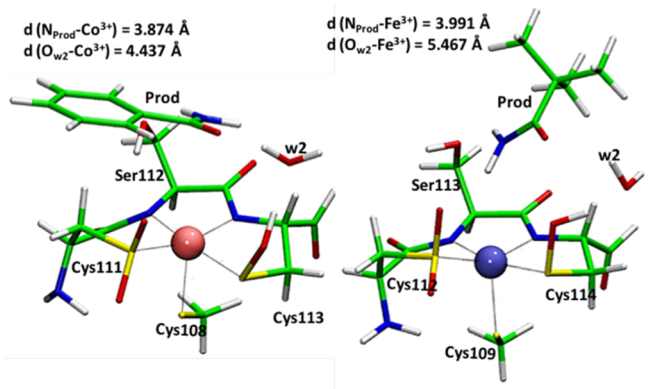


Figure 5. M06*L*/6-31+G(d,p) (SDD for metal ion) optimized structures of the EP complex for Co- and Fe-types. For clarity, only the amino acid residues of the inner coordination shell of the metal centers are retained.

negative when the examined distance decreases. In this situation, it is clear that release of the product is a barrierless process. The water molecule has completely replaced the product and is coordinated with cobalt ion at a distance of 2.246 Å, whereas the product (P) which now lies far from the metal coordination sphere (4.745 Å) is held in the catalytic pocket through weak interactions involving the Leu48 and Tyr68 residues that make up the walls of the pocket. The relative energy values of the complex E + P is −14.9 kcal/mol (see Figure 2).

Fe−NHase. In last years, different theoretical and experimental works devoted to studying in particular Fe-type nitrile hydratases, appeared in literature.[27,29,30,56−58] Yamanaka et al.[54] investigated the reaction mechanism of an Fe-type NHase from *Rhodococcus erythropolis* N771 (ReNHase) using time-resolved X-ray crystallography and a *tert*-butylnitrile or pivalonitrile (PivCN) as substrate. In this work, authors proposed that the metal-coordinated substrate is nucleophilically attacked by the O(SO^-) atom of αCys114−SO^-, and that this step is followed by another nucleophilic attack to the S(SO^-) atom by the βArg56-activated water molecule to release the amide product and regenerate αCys114−SO^-.[54] While taking due account of the crystallographic study by Yamanaka and co-workers,[54] our investigation presents some innovative aspect. In particular, this concerns the role of the two water molecules which, as will see, will play an active role in the mechanism. In the optimized structure of enzyme without substrate (E) these two water molecules are linked by a H-bond (1.784 Å) and are located at 2.128 and 4.034 Å from the iron. Besides, one of them forms a H-bond with the −OH of the sulfenic group (1.585 Å). In any case, the mechanism of the iron containing nitrile hydratase is the same of the cobalt containing one. This is also affirmed by Yamanaka et al.,[54] who exclude the formation of a disulfide intermediate and support that of cyclic intermediate just like in the Co-type enzyme.[17] The energetic profile for the hydrolysis of pivalonitrile by Fe(III)−NHase is reported in Figure 2. As in the case of cobalt containing enzyme, preliminary computations on the ES, have confirmed that the lowest-energy spin state for the metal ion in Fe(III)−NHase is S = 1/2 in agreement with experimental evidence.[25]

The displacement of water molecule from the axial position by pivalonitrile generates the Michaelis−Menten complex (ES) whose optimized geometry is depicted in Figure 3. Upon

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

coordination to the low spin Fe^{3+} cation, the nitrile is activated. A σ donation to the metal from the lone pair of nitrile nitrogen depletes the electron density on the ligand donor atom. This means that nitrogen tends to become less negative. As consequence, an electron attraction occurs from the adjacent carbon which becomes more electrophile. The trend of charge distributions for both enzymes (see Figure 6) depicts this behavior proposing a carbon nitrile more positive (by about +0.191 e) in the ES, with respect to that of the free substrates.

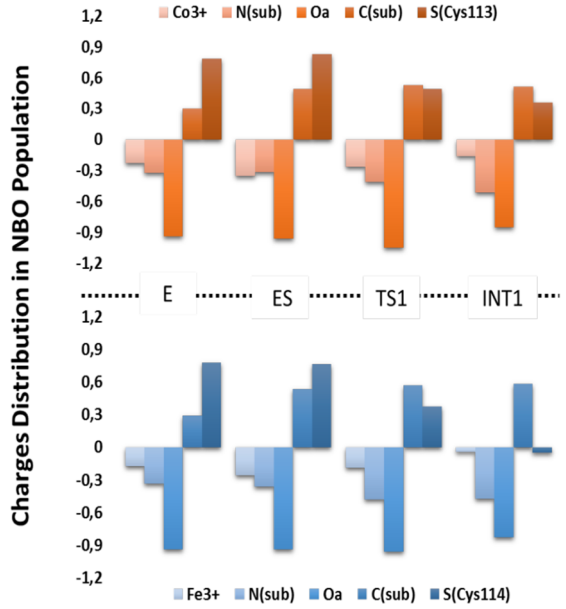


Figure 6. Charge distributions in some significant stationary points on the PES of Co(III)− and Fe(III)−NHase enzymes.

In the ES complex, water (**w1**) lies at 4.357 Å from the metal center, whereas the nucleophile agent (S−OH) is at 3.949 Å from the nitrile carbon (distance in the Co-type enzyme was 3.533 Å). The nucleophilic attack of cysteine−sulfenic acid on the substrate requires 16.2 kcal/mol (TS1). The corresponding optimized structure is shown in Figure 7. The enzyme reducing the barrier for nitrile hydration, produces a catalytic effect of around 20 kcal/mol in comparison with the uncatalyzed reaction (35.6 kcal/mol at B3LYP level).[56] This is not the alone significant result. In fact, considering the protonated αCys114−SOH as new nucleophile agent as suggested by the experimental observation,[28,29] which is different from those previously proposed,[27,29,57,59] it is possible to assess that the used model assures a good description of the interactions involved in the nucleophilic attack step. Moreover, it is important to underline the role of the βArg56 that assists this step as never before in the previous theoretical works.[56−60] The water molecule close to the S atom of αCys114−SOH forms a hydrogen bond with the O atom of this residue 1.780 Å. This behavior is in agreement with the experimental observation.[54] The TS1 (see Figure 7) assumes a cyclic structure with the S−OH bond just broken (2.894 Å) and the C−OH that is forming (1.786 Å). The angle (137.9°) between the carbon atom of the pivaloyl group and the C and N atoms of the nitrile group is no longer linear. In addition, these three atoms and the sulfenic (S−OH) group of the modified αCys114 lie in a plane. The reorganization involving the sulfenic group is assisted by the βArg56 residue that is implicated in a H-bond with its OH moiety. The different values of ionic radius (LS Co^{3+}=53 pm, LS Fe^{3+} = 55 pm), the length of the Csp−OH bond (2.015 Å in Co−NHase and 1.786 Å in Fe−NHase) in TS1, and the bond orders (0.239 and 0.305) arising from atom−atom overlap-weighted NAO analysis for Co and Fe ions indicate the more advanced nucleophilic attack in the Fe−NHase, which finds confirmation in the lower activation barrier. A further reasonable basis for evaluating the different energetic behavior of the two Co- and Fe-dependent nitrile hydratases can derive from MO energy diagram in Figure 8. Here it is possible to observe that the energy of the HOMO β and LUMO β in the ES complex of Fe(III)−NHase decreases much more than that of the same orbitals in the Co(III)−NHase. This means that a better overlap occurs between the HOMO β of free enzyme with the HOMO of nitrile with a consequent stronger covalent interaction. The intermediate (INT1) confirms the elongation of α Cys114 OH---S distance (3.242 Å). The C−OH bond is now formed (1.359 Å). The plane identified by N═C−OH−S moiety deviates from planarity by 24.4°. The INT1 proceeds toward the final product through TS2 lying at 14.2 kcal/mol above reactants. Unlike what has been proposed in previous works where the deprotonated form of the sulfenic group was used, TS2 describes the proton's delivery to the nitrogen by the −OH group of the iminol group. This occurs through a concerted proton transfer mediated by the cluster of the two water molecules inserted between the −OH of the iminol group and the −OH of the serine (1.304 Å). In fact, the **w1** receives the proton (1.371 Å) from the N═C−OH moiety, giving the other one to **w2** (1.389 Å) that through the serine delivers it to the nitrogen atom of the forming amidate. The INT2 lies at 24.1 above the ES and shows the consequence of the previously described proton transfer (see Figure S1).

As found in the Co−NHase, after INT1 the formation of a disulfide bond is observed even if in the last step the CH_3−S− moiety fits again in an axial position. In any case, this is dependent neither on the used QM approach nor on the following mechanism because in other previous theoretical works the S−S bond is also observed.[29,57] The reaction evolves to the final product through the TS3. As previously found for Co−NHase, this stage is related to the formation of the S−O bond of αCys114 at the expense of one water molecule. In fact, the Arg56 residue ensures the position of the **w1** and **w2** molecules next to the site where the reaction takes place, through hydrogen bond network. The H_2O approaches to the sulfur atom of αCys114 for attacking it (2.466 Å) and simultaneously gives a proton (1.351 Å) to the −NH moiety of the deprotonated amide still coordinated to the iron center 2.103 Å (Figure 7). The reported imaginary frequency well accounts for the concerted stretching motions of the S−O and H−NH bonds. The energetic request for this is about 8 kcal/mol. It is worth noting that what occurs in this phase is in agreement with the findings of the experimental counterpart based on FTIR analyses of NHase.[54] In fact, following hydration in $H_2{}^{18}O$, Yamanaka et al.[54] note that the oxygen atom of the αCys114 residue is just the one deriving from water (^{18}O). In the EP species, the amide formation is completed but the initial active site structure is not totally regenerated. The distance of the amide nitrogen from iron of 3.991 Å suggests that the product is released. The **w2** molecule is engaged in a network of hydrogen bond involving βArg56, αCys114, and αGln90 residues. EP lies at about 23.7 kcal/mol below the separated reactants.

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

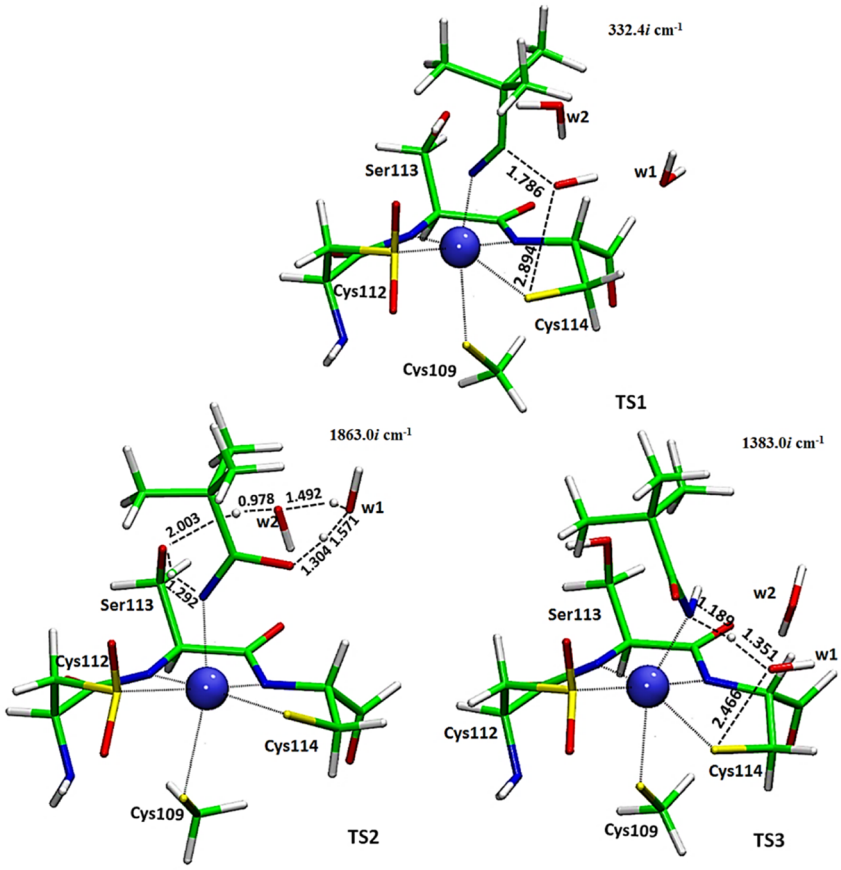


Figure 7. M06*L*/6-31+G(d,p) (SDD for metal ion) optimized structures of the transition states for Co-NH type. For clarity, only the amino acid residues of the inner coordination shell of the metal center are retained. For each TS, imaginary frequencies values are reported.

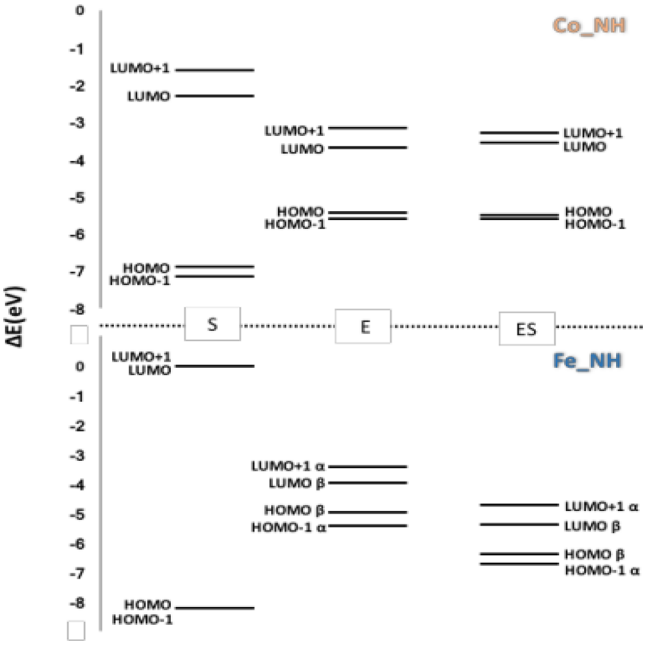


Figure 8. Energy diagram of frontier molecular orbitals of nitrile substrates (S), free enzymes (E), and substrate complexes (ES).

A look at the energetics of the two enzymatic reactions examined shows that the highest activation barriers occur in both cases for the formation of the EP complex, that is, the one that we have indicated as the phase of restoring of the catalyst. In the case of cobalt, this barrier is very similar to that of the first nucleophile attack (the difference is less than 1 kcal/mol). Instead, in the case of iron, the values are very different. However, since the formation of the amide has already practically occurred in correspondence with the INT2, we could hypothesize that the nucleophilic attack by the − OH group of the αCys113−S−OH on the nitrile carbon atom is the lowest step of both process. Moreover, at least in the case of iron, the experimental data concerning the activation barrier are very controversial, and thus we cannot say with certainty that our assumption is definitive. The comforting data are that in any case the barriers to be overcome are not very high and are compatible with the catalytic processes.

■ CONCLUSION

The present work reports the results of a fully QM study performed in the framework of density functional theory and using the cluster approach. It is devoted to the comparison of the catalytic efficiency of the two Fe(III)−NHase and Co(III)−NHase enzymes toward the pivalonitrile and benzonitrile substrates, respectively. The models used for computations were built up cutting out the metalloprotein's structures reported in Data Bank deriving from the *Rhodococcus erythropolis* (pdb code: 2ZPE) and *Pseudonocardia Thermophila* (pdb code: 1IRE) microorganisms. These present some difference in the amino acid sequence of the outer coordination sphere around the metal ion. Consequently, two clusters of different size have been obtained. The explored mechanism takes into account the most recent experimental observations

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

on these two enzymes, that propose the sulfenic group of the modified cysteine coordinated to the metal ion as nucleophile agent. Furthermore, our theoretical study suggests in both examined cases the two water molecules present in the active site as indirect assistants during the last part of catalytic reaction. Such a role was never explored before. Because of the outcomes of our investigation, it is possible to draw the following conclusions:

- The spin state of Fe^{3+} and Co^{3+} ions in enzyme−substrate complexes was confirmed by preliminary calculations to be 1/2 and 0, respectively, in agreement with experimental indication on nitrile hydratases.
- The examined mechanism seems to be plausible for both enzymes in that the energy requests fall within the typical values for catalytic events.
- The enzyme restoring phase described by TS2 and TS3 transition states represents the most expensive phase of catalytic cycle from the energetic point of view for both iron and cobalt containing enzymes in agreement with experimental findings.
- The values of activation barriers suggest that the iron-dependent enzyme works better than Co one but the catalytic power of the Co-type is not compromised. A justification of this different behavior was found into the energetic distribution of the frontier orbitals during the ES complex formation that suggest a stronger covalent interaction between the HOMO β of free enzyme with the HOMO of nitrile substrate in the case of iron containing enzyme that favors the next nucleophilic attack.
- The crystallographic water molecules retained in the cluster of Fe(III)− and Co(III)−NHases play an active role in the catalytic process. In particular, **w1** and **w2** mediate the multiple proton transfer between the αCys113−OH and the nitrogen atom of nitrile that occurs in TS2 and contribute to the stabilization of the EP complex. Beside, **w1** directly comes into play in TS3, that is, the nucleophilic attack on the S atom of αCys 113,114, for establishing the nucleophile agent αCys113(114)−S−OH necessary for the catalytic cycle restore. Such behaviors are consistent with the experimental observations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.inorgchem.7b02121.

Additional optimized structures and results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: tiziana.marino65@unical.it (T.M.).
*E-mail: m.toscano@unical.it (M.T.).

ORCID

Tiziana Marino: 0000-0003-2386-9078

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support from the Università degli Studi della Calabria, Dipartimento di Chimica e Tecnologie Chimiche (CTC), is acknowledged.

■ REFERENCES

(1) Vennesland, B. E. E.; Conn, C. J.; Knowles, J.; Westley, J.; Wissing, F. *Cyanide in Biology*; Academic Press, London, 1981.
(2) Conn, E. E. *Cyanide Compounds in Biology*; Wiley: Chichester, U.K., 1988.
(3) Takahashi, N. *Chemistry of Plant Hormones*; CRC Press: Boca Raton, FL, 1986.
(4) Gibbs, R. D. *Chemotaxonomy of Flowering Plants*; McGill-Queens University Press: Montreal, Ontario, 1974.
(5) Mascharak, P. K. Structural and functional models of nitrile hydratase. *Coord. Chem. Rev.* **2002**, *225*, 201−214.
(6) Leonard, A.; Gerber, G. B.; Stecca, C.; Rueff, J.; Borba, H.; Farmer, P. B.; Sram, R. J.; Czeizel, A. E.; Kalina, I. Mutagenicity, carcinogenicity, and teratogenicity of acrylonitrile. *Mutat. Res., Rev. Mutat. Res.* **1999**, *436*, 263−283.
(7) Wyatt, J. M.; Knowles, C. J. The development for microbial treatment of acrylonitrile effluents. *Biodegradation* **1995**, *6*, 93−107.
(8) Woutersen, R. A. Psychosocial factors at work and subsequent depressive symptoms in the Gazel cohort. *Scand. J. Work. Environ. Health* **1998**, *24* (5), 197−205.
(9) Delzell, E.; Monson, R. R. Mortality among rubber workers. VI. Men with exposure to Acrylonitrile. *J. Occup. Med.* **1982**, *24*, 767−770.
(10) Czeizel, A. E.; Hegedus, S.; Timar, L. Congenital abnormalities and indicators of germinal mutations in the vicinity of an acrylonitrile producing factory. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **1999**, *427*, 105−123.
(11) Kidd, H.; James, D. R. *The Agrochemicals Handbook*, 3rd ed.; Royal Society of Chemical Information Services: Cambridge, UK, 1998.
(12) Wyatt, J. M.; Linton, E. A. The industrial potencial of microbial nitrile biochemistry. In *Cyanide Compounds in Biology*;Evered, D., Harnett, S., Eds.; Ciba Foundation Symposium; Wiley−Interscience: Hoboken, NJ, 1998; Vol. *140*, pp 32−48.
(13) Yamada, H.; Shimizu, S.; Kobayashi, M. Hydratases involved in nitrile conversion: Screening, characterization and application. *Chem. Rec.* **2001**, *1*, 152−161.
(14) Novak, W.; Ohtsuka, Y.; Hasegawa, J.; Nakatsuji, H. Density functional study on geometry and electronic structure of nitrile hydratase active site model. *Int. J. Quantum Chem.* **2002**, *90*, 1174−1187.
(15) *Hazardous Substances Databank*; U.S. National Library of Medicine: Bethesda, MD, 1995.
(16) Wyatt, J. M.; Knowles, C. J. Microbial degradation of acrylonitrile waste effluents: the degradation of effluents and condensates from the manufacture of acrylonitrile. *Int. Biodeterior. Biodegrad.* **1995**, *35* (1-3), 227−248.
(17) Martinez, S.; Wu, R.; Sanishvili, R.; Liu, D.; Holz, R. The Active Site Sulfenic Acid Ligand in Nitrile Hydratases Can Function as a Nucleophile. *J. Am. Chem. Soc.* **2014**, *136*, 1186−1189.
(18) Kobayashi, M.; Shimizu, S. Metalloenzyme NHase: Structure, regulation and application to biotechnology. *Nat. Biotechnol.* **1998**, *16*, 733−736.
(19) Baxter, J.; Cummings, S. *Antonie van Leeuwenhoek* **2006**, *90*, 1−17.
(20) Asano, Y.; Tani, Y.; Yamada, H. A new enzyme ″nitrile hydratase″ which degrades acetonitrile in combination with amidase. *Agric. Biol. Chem.* **1980**, *44*, 2251−2252.
(21) Ashima, Y.; Suto, M. Development of an Enzymatic Process for Manufacturing Acrylamide and Recent Progress. *Bioprocess Technol.* **1993**, *16*, 91−107.

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

(22) Kobayashi, M.; Nagasawa, T.; Yamada, H. Enzymatic synthesis of acrylamide: a success story not yet over. *Trends Biotechnol.* **1992**, *10*, 402−408.
(23) Yamada, H.; Kobayashi, M. Nitrile Hydratase and Its Application to Industrial Production of Acrylamide. *Biosci., Biotechnol., Biochem.* **1996**, *60*, 1391−1400.
(24) Payne, M. S.; Wu, S.; Fallon, R. D.; Tudor, G.; Stieglitz, B.; Turner, I. M., Jr.; Nelson, M. J. A Stereoselective Cobalt-Containing Nitrile Hydratase. *Biochemistry* **1997**, *36* (18), 5447−5454.
(25) Kovacs, J. A. Synthetic Analogues of Cysteinate-Ligated Non-Heme Iron and Non-Corrinoid Cobalt Enzymes. *Chem. Rev.* **2004**, *104*, 825−848.
(26) Hopmann, K. H.; Himo, F. Theoretical Investigation of the Second-Shell Mechanism of Nitrile Hydratase. *Eur. J. Inorg. Chem.* **2008**, *20*, 1406−1412.
(27) Dey, A.; Chow, M.; Taniguchi, K.; Lugo-Mas, P.; Davin, S.; Maeda, M.; Kovacs, J. A.; Odaka, M.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Sulfur K-Edge XAS and DFT Calculations on Nitrile Hydratase: Geometric and Electronic Structure of the Non-heme Iron Active Site. *J. Am. Chem. Soc.* **2006**, *128*, 533−541.
(28) Noguchi, T.; Nojiri, M.; Takei, K.; Odaka, M.; Kamiya, N. Protonation Structures of Cys-Sulfinic and Cys-Sulfenic Acids in the Photosensitive Nitrile Hydratase Revealed by Fourier Transform Infrared Spectroscopy. *Biochemistry* **2003**, *42*, 11642−11650.
(29) Light, K. M.; Yamanaka, Y.; Odaka, M.; Solomon, E. I. Spectroscopic and computational studies of nitrile hydratase: insights into geometric and electronic structure and the mechanism of amide synthesis. *Chem. Sci.* **2015**, *6*, 6280−6294.
(30) Stein, N.; Gumataotao, N.; Hajnas, N.; Wu, R.; Lankathilaka, K. P. W.; Bornscheuer, U. T.; Liu, D.; Fiedler, A. T.; Holz, R. C.; Bennett, B. Multiple States of Nitrile Hydratase from Rhodococcus equi TG328−2: Structural and Mechanistic Insights from Electron Paramagnetic Resonance and Density Functional Theory Studies. *Biochemistry* **2017**, *56*, 3068−3077.
(31) Banerjee, A.; Sharma, R.; Banerjee, U. The nitrile-degrading enzymes: current status and future prospects. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 33−44.
(32) Mylerova, V.; Martinkova, L. Synthetic applications of nitrile-converting enzymes. *Curr. Org. Chem.* **2003**, *7*, 1279−1295.
(33) Amata, O.; Marino, T.; Russo, N.; Toscano, M. Catalytic activity of a z-class zinc and cadmium containing carbonic anhydrase. Compared work mechanisms. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3468−3477.
(34) Amata, O.; Marino, T.; Russo, N.; Toscano, M. A Proposal for Mitochondrial Processing Peptidase Catalytic Mechanism. *J. Am. Chem. Soc.* **2011**, *133* (44), 17824−17831.
(35) Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780−6786.
(36) Marino, T.; Prejanò, M.; Russo, N. How Can Methanol Dehydrogenase from Methylacidiphilum fumariolicum Work with the Alien CeIII Ion in the Active Center? A Theoretical Study. *Chem. - Eur. J.* **2017**, *23*, 8652−8657.
(37) Miyanaga, A.; Fushinobu, S.; Ito, K.; Wakagi, T. Crystal structure of cobalt-containing nitrile hydratase. *Biochem. Biophys. Res. Commun.* **2001**, *288*, 1169−1174.
(38) Hashimoto, K.; Suzuki, H.; Taniguchi, K.; Noguchi, T.; Yohda, M.; Odaka, M. Catalytic Mechanism of Nitrile Hydratase Proposed by Time-resolved X-ray Crystallography Using a Novel Substrate, tert-Butylisonitrile. *J. Biol. Chem.* **2008**, *283*, 36617−36623.
(39) Martinez, S.; Yang, X.; Bennett, B.; Holz, R. C. A cobalt-containing eukaryotic nitrile hydratase. *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865*, 107−112.
(40) Zhou, Z.; Hashimoto, Y.; Cui, T.; Washizawa, Y.; Mino, H.; Kobayashi, M. Unique Biogenesis of High-Molecular Mass Multimeric Metalloenzyme Nitrile Hydratase: Intermediates and a Proposed Mechanism for Self-Subunit Swapping Maturation. *Biochemistry* **2010**, *49*, 9638−9648.
(41) Martinez, S.; Wu, R.; Sanishvili, R.; Liu, D.; Holz, R. The Active Site Sulfenic Acid Ligand in Nitrile Hydratases Can Function as a Nucleophile. *J. Am. Chem. Soc.* **2014**, *136*, 1186−1189.
(42) Siegbahn, P. E. M.; Blomberg, M. R. A. Transition-Metal Systems in Biochemistry Studied by High-Accuracy Quantum Chemical Methods. *Chem. Rev.* **2000**, *100*, 421−443.
(43) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601−3658 and references therein..
(44) Jçnsson, T. J.; Tsang, A. W.; Lowther, C. M.; Furdui, W. T. Identification of Intact Protein Thiosulfinate Intermediate in the Reduction of Cysteine Sulfinic Acid in Peroxiredoxin by Human Sulfiredoxin. *J. Biol. Chem.* **2008**, *283*, 22890−22894.
(45) Nagashima, S.; Nakasako, M.; Dohmae, N.; Tsujimura, M.; Takio, K.; Odaka, M.; Yohda, M.; Kamiya, N.; Endo, I. Novel non-heme iron center of nitrile hydratase with a claw setting of oxygen atoms. *Nat. Struct. Biol.* **1998**, *5*, 347−351.
(46) Frisch, M. J.; G. W., Trucks, Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2013.
(47) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.
(48) Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. Energy-Adjusted Ab Initio Pseudopotentials for the Second and Third Row Transition Elements. *Theor. Chim. Acta* **1990**, *77*, 123−141.
(49) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. *Natural Bond Orbital NBO*, version 3.1; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI.
(50) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(51) Warshel, A. *In Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley, New York, 1991.
(52) Ramos, M. J.; Fernandes, P. A. Computational Enzymatic Catalysis. *Acc. Chem. Res.* **2008**, *41*, 689−698.
(53) Piazzetta, P.; Marino, T.; Russo, N. The working mechanism of the β-carbonic anhydrase degrading carbonyl sulphide (COSase): a theoretical study. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14843−14848.
(54) Yamanaka, Y.; Kato, Y.; Hashimoto, K.; Iida, K.; Nagasawa, K.; Nakayama, H.; Dohmae, N.; Noguchi, K.; Noguchi, T.; Yohda, M.; Odaka, M. Time-Resolved Crystallography of the Reaction Intermediate of Nitrile Hydratase:Revealing a Role for the Cysteine sulfenic Acid Ligand as a Catalytic Nucleophile. *Angew. Chem., Int. Ed.* **2015**, *54*, 10763−10767.
(55) Hopmann, K. H.; Guo, J.-D.; Himo, F. Theoretical Investigation of the First-Shell Mechanism of Nitrile Hydratase. *Inorg. Chem.* **2007**, *46*, 4850−4856.
(56) Hopmann, K. H.; Himo, F. On the Role of Tyrosine as Catalytic Base in Nitrile Hydratase. *Eur. J. Inorg. Chem.* **2008**, *22*, 3452−3459.
(57) Hopmann, K. H. Full Reaction Mechanism of Nitrile Hydratase: A Cyclic Intermediate and an Unexpected Disulfide Switch. *Inorg. Chem.* **2014**, *53*, 2760−2762.

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

(58) Kayanuma, M.; Hanaoka, K.; Shoji, M.; Shigeta, Y. A QM/MM study of the initial steps of catalytic mechanism of nitrile hydratase. *Chem. Phys. Lett.* **2015**, *623*, 8−13.
(59) Kayanuma, M.; Shoji, M.; Yohda, M.; Odaka, M.; Shigeta, Y. Catalytic Mechanism of Nitrile Hydratase Subsequent to Cyclic Intermediate Formation: A QM/MM Study. *J. Phys. Chem. B* **2016**, *120*, 3259−3266.
(60) Takarada, H.; Kawano, Y.; Hashimoto, K.; Nakayama, H.; Ueda, S.; Yohda, M.; Kamiya, N.; Dohmae, N.; Maeda, M.; Odaka, M. Mutational study on αGln90 of Fe-type nitrile hydratase from Rhodococcus sp. N771. *Biosci., Biotechnol., Biochem.* **2006**, *70*, 881−889.

DOI: 10.1021/acs.inorgchem.7b02121
Inorg. Chem. 2017, 56, 13390−13400

Reaction Mechanism of low spin Iron (III)- and Cobalt (III) containing Nitrile Hydratases: A Quantum Mechanics Investigation

Tiziana Marino*, Mario Prejanò, Carmen Rizzuto, Josè Carlos Madrid Madrid, Nino Russo, Marirosa Toscano*

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)

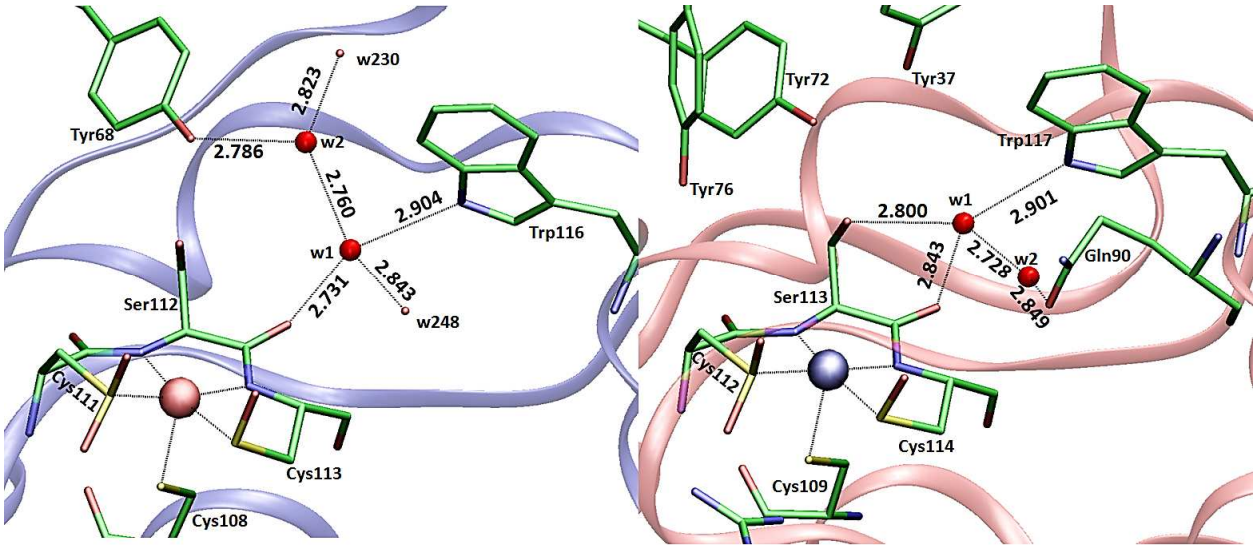


Figure S1. X ray hydrogen bonds network for water molecules explicitly considered in the model for the Co-NHase and Fe-NHase, on the left and on the right respectively.

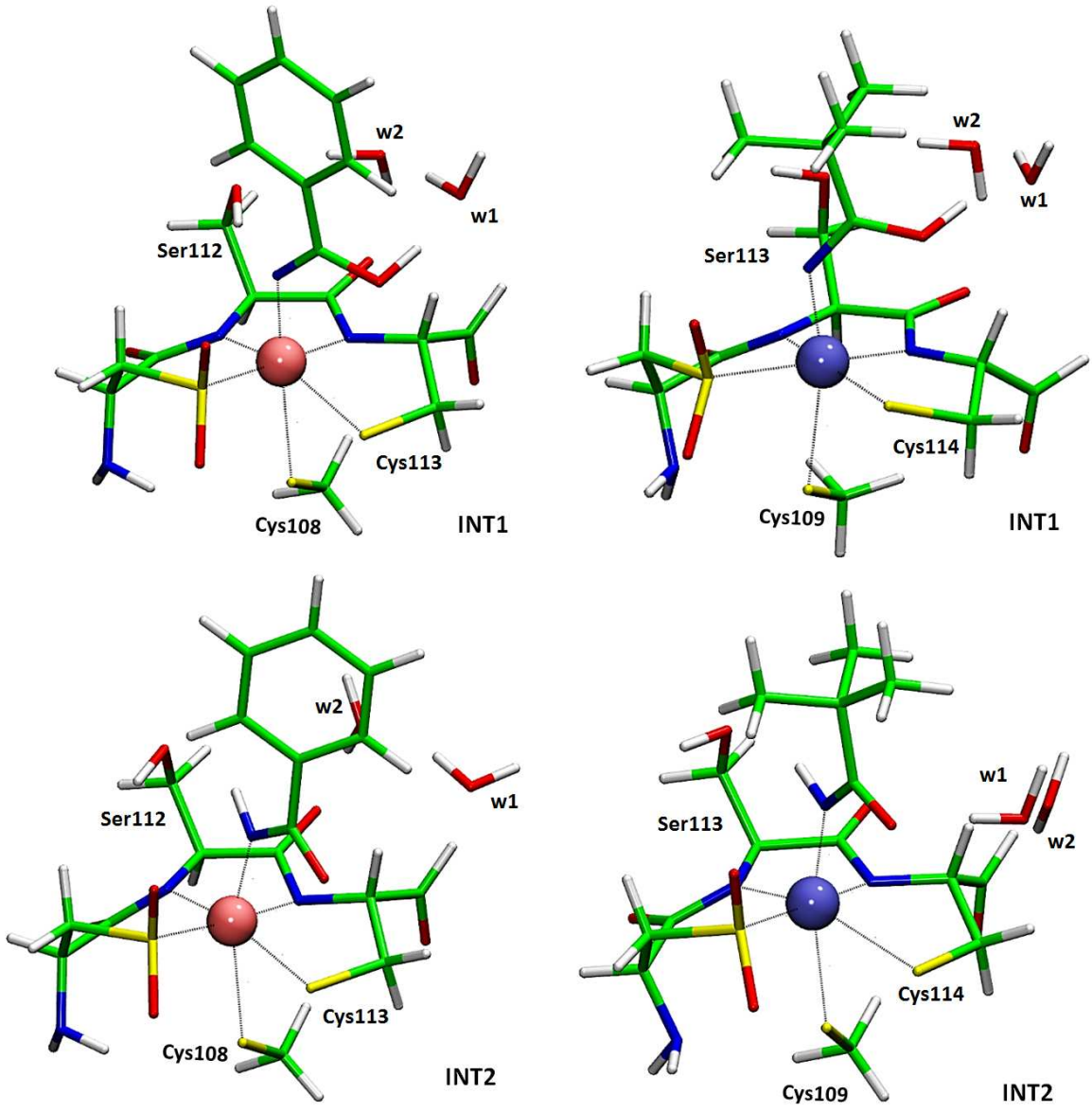


Figure S2. Optimized M06L structures of the two intermediates intercepted along potential energy surfaces for the Co-NHase and Fe-NHase. For clarity, only the amino acid residues of the inner coordination shell of the metal centre are retained

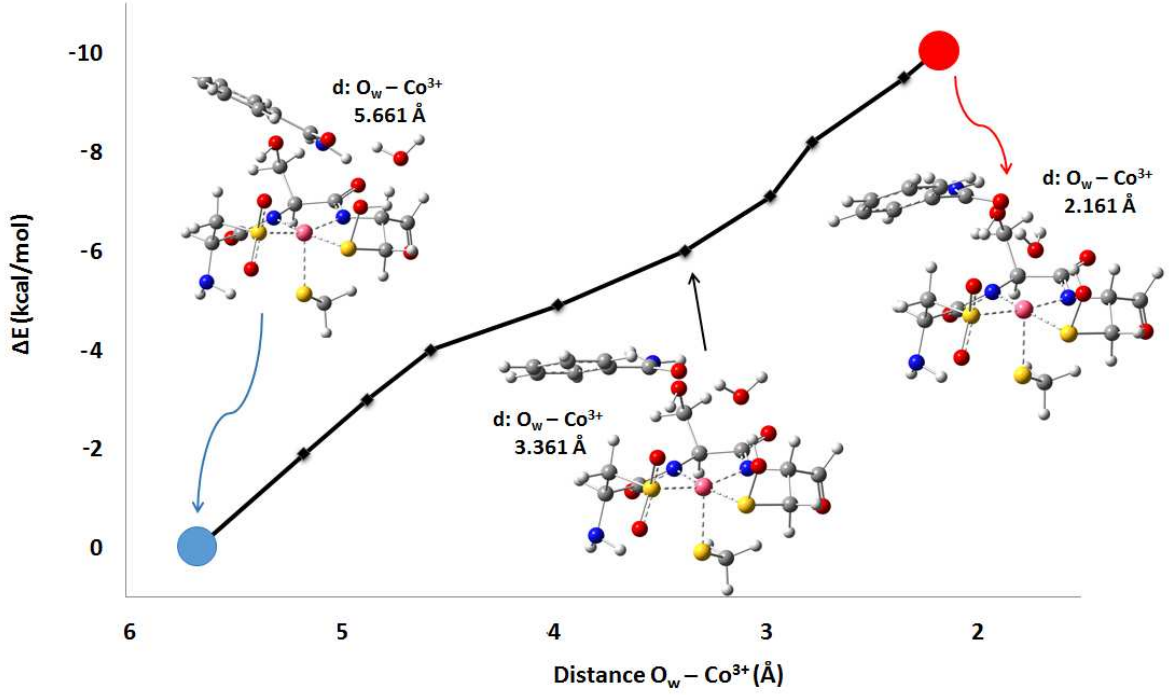


Figure S3. Potential energy scan for the approach of H_2O (**w1**) molecule to the EP complex of Co-NHase as a function of H_2O-Co distance.

Table S1. Cartesian coordinates for E CoNH

C	1.41200000	-4.99100000	-0.27500000
S	0.02500000	-3.92400000	-0.73600000
N	0.88400000	-2.35700000	-3.87200000
C	1.53800000	-1.34600000	-3.04600000
C	0.56400000	-0.26700000	-2.60900000
S	-0.66000000	-0.86500000	-1.43500000
C	2.44800000	-1.92400000	-1.95000000
O	3.45000000	-2.53200000	-2.34200000
O	-1.34900000	0.39000000	-0.90800000
O	-1.61500000	-1.71600000	-2.25900000
N	2.18700000	-1.66900000	-0.62500000
C	3.27300000	-2.13400000	0.25700000
C	2.74200000	-2.44200000	1.63900000
O	3.47700000	-2.58300000	2.62800000
C	4.35700000	-1.07300000	0.32300000
O	3.78200000	0.18200000	0.72100000
N	1.40800000	-2.56500000	1.66900000
C	0.76300000	-2.83700000	2.94300000
C	-0.56900000	-3.48900000	2.62600000
S	-1.43100000	-2.56400000	1.31000000
C	1.61000000	-3.79500000	3.79800000
O	1.66700000	-4.97800000	3.54400000
O	2.25600000	0.47900000	3.01900000
C	-6.15100000	5.04500000	3.70700000
C	-5.78000000	4.19400000	2.49800000
C	-6.02600000	4.95200000	1.19800000
C	-4.33500000	3.71700000	2.58400000
C	-4.14400000	1.68100000	-1.07500000
N	-3.89600000	0.75400000	0.00800000
C	-4.79300000	0.48600000	0.94700000
N	-6.02200000	1.05300000	0.86400000
N	-4.53300000	-0.39700000	1.91400000
C	9.61600000	3.76200000	-1.77900000
C	8.39400000	3.45500000	-0.96700000
C	7.58500000	2.35900000	-1.27000000
C	8.03300000	4.23900000	0.13700000
C	6.46400000	2.04400000	-0.50500000
C	6.91900000	3.94000000	0.91300000
C	6.12700000	2.83300000	0.59700000
O	5.04800000	2.56900000	1.38400000
C	-6.92100000	-1.58800000	-3.03600000
N	-6.03200000	-2.38800000	-2.21800000
C	-4.69800000	-2.21300000	-2.21600000
N	-4.15900000	-1.24200000	-2.96000000
N	-3.91200000	-2.95600000	-1.44100000
Co	0.42700000	-1.91500000	0.19000000
O	0.29800000	-0.05300000	1.32300000
O	-1.92300000	-1.20300000	2.26200000
H	5.09400000	-1.33300000	1.08600000
H	4.85000000	-0.97400000	-0.65000000
H	3.02800000	0.34600000	0.12900000
H	3.73500000	-3.04900000	-0.14000000
H	-5.53100000	5.94800000	3.74700000
H	-6.00000000	4.50200000	4.64500000
H	-6.43600000	3.30700000	2.50500000

H	-5.43200000	5.87300000	1.17000000
H	-5.73800000	4.36300000	0.31500000
H	-7.07500000	5.24100000	1.08400000
H	-3.65000000	4.57200000	2.64900000
H	-4.03700000	3.13900000	1.69800000
H	-4.15900000	3.09600000	3.46900000
H	-4.99800000	1.36900000	-1.68500000
H	-3.24800000	1.69200000	-1.69600000
H	-2.95600000	0.33500000	0.01100000
H	-5.16500000	-0.46800000	2.69400000
H	-3.56300000	-0.72200000	2.06000000
H	-6.60900000	1.02700000	1.68200000
H	-6.10900000	1.90300000	0.32600000
H	10.53400000	3.49100000	-1.24600000
H	9.61700000	3.21200000	-2.72400000
H	7.83500000	1.73200000	-2.12400000
H	8.64100000	5.10300000	0.39900000
H	5.84800000	1.18600000	-0.76200000
H	6.65200000	4.55000000	1.77000000
H	4.64200000	1.73800000	1.07700000
H	-6.76800000	-0.52400000	-2.83500000
H	-6.79000000	-1.77800000	-4.10800000
H	-6.39100000	-3.24300000	-1.82200000
H	-4.65300000	-0.92600000	-3.77900000
H	-3.12900000	-1.17900000	-2.94100000
H	-4.28900000	-3.58100000	-0.74900000
H	-2.89800000	-2.81300000	-1.50600000
H	1.53200000	-5.07000000	0.80800000
H	2.33400000	-4.62200000	-0.73300000
H	0.27900000	-2.92900000	-3.28500000
H	1.61100000	-2.98100000	-4.21300000
H	2.25000000	-0.83100000	-3.70400000
H	1.08000000	0.54800000	-2.09100000
H	0.00100000	0.12300000	-3.46300000
H	0.61500000	-1.90300000	3.51000000
H	-1.25300000	-3.54400000	3.47700000
H	-0.40700000	-4.49600000	2.22600000
H	2.09100000	-3.34400000	4.68600000
H	1.05700000	0.18400000	1.93300000
H	0.19500000	0.68600000	0.70100000
H	-1.18600000	-0.55500000	2.13100000
H	3.03200000	0.25700000	2.46900000
H	-4.30800000	2.70300000	-0.70500000
H	-7.94900000	-1.83200000	-2.77600000
H	9.69200000	4.82800000	-2.01500000
H	1.19800000	-5.98100000	-0.68700000
H	-7.19600000	5.36600000	3.67000000
H	2.43000000	1.37000000	3.33800000

Table S2. Cartesian coordinates for ES CoNH

C	0.56400000	-4.94000000	1.05900000
S	-0.63900000	-3.87500000	0.18000000
N	0.72700000	-3.60400000	-3.22100000
C	1.42600000	-2.40200000	-2.76000000
C	0.51900000	-1.16600000	-2.79700000
S	-0.86100000	-1.23400000	-1.62400000
C	2.25400000	-2.61900000	-1.47000000
O	3.29400000	-3.28000000	-1.60200000
O	-1.50100000	0.15100000	-1.65300000

O	-1.82700000	-2.27100000	-2.20300000
N	1.88000000	-2.05200000	-0.28600000
C	2.86600000	-2.23900000	0.80400000
C	2.19500000	-2.03500000	2.15200000
O	2.85100000	-1.87500000	3.20800000
C	4.01400000	-1.23800000	0.64000000
O	3.44400000	0.08400000	0.45600000
N	0.86800000	-2.03300000	2.09900000
C	0.13400000	-1.79900000	3.33800000
C	-1.29400000	-2.28900000	3.12100000
S	-1.89900000	-1.64100000	1.51300000
C	0.76200000	-2.53200000	4.54000000
O	0.66800000	-3.73200000	4.68200000
O	-2.24100000	-0.03400000	2.00500000
C	-5.71500000	6.65200000	1.01900000
C	-5.33900000	5.17000000	0.88000000
C	-3.97400000	5.00100000	0.19600000
C	-5.35600000	4.47100000	2.24900000
C	-3.98000000	1.71000000	-2.26400000
N	-3.89000000	1.18600000	-0.90800000
C	-4.84400000	1.32100000	0.00500000
N	-6.01900000	1.92000000	-0.32700000
N	-4.69000000	0.82700000	1.23800000
C	9.95300000	1.79500000	-2.33700000
C	8.56900000	1.97100000	-1.75600000
C	7.63500000	0.92700000	-1.78700000
C	8.18500000	3.16900000	-1.13400000
C	6.37000000	1.06400000	-1.21100000
C	6.92900000	3.32000000	-0.54600000
C	6.01700000	2.26000000	-0.57400000
O	4.79300000	2.43200000	0.02200000
C	-7.09700000	-1.66100000	-3.34500000
N	-6.31300000	-2.34700000	-2.31900000
C	-4.96400000	-2.33300000	-2.26900000
N	-4.26200000	-1.67500000	-3.20200000
N	-4.31000000	-2.92800000	-1.26900000
Co	0.07600000	-1.73300000	0.39500000
O	-0.31500000	1.52200000	3.03900000
O	2.30700000	0.93600000	3.00100000
H	4.62700000	-1.20300000	1.54200000
H	4.62300000	-1.50400000	-0.22700000
H	2.69500000	-0.07000000	-0.15000000
H	3.29600000	-3.24800000	0.78300000
H	-4.98600000	7.17600000	1.64800000
H	-6.70100000	6.77200000	1.48000000
H	-6.10100000	4.69500000	0.24100000
H	-3.18100000	5.47200000	0.78800000
H	-3.70200000	3.94400000	0.08200000
H	-3.96600000	5.46600000	-0.79700000
H	-4.64200000	4.94800000	2.93100000
H	-5.06300000	3.41600000	2.17000000
H	-6.34500000	4.52700000	2.71800000
H	-4.82200000	1.26800000	-2.81000000
H	-3.05700000	1.43300000	-2.77000000
H	-3.03000000	0.64800000	-0.72700000
H	-5.41500000	0.96100000	1.92600000
H	-3.76000000	0.51500000	1.56900000
H	-6.59900000	2.27500000	0.41900000
H	-6.06500000	2.45700000	-1.17900000
H	10.66800000	1.47800000	-1.56700000

H	9.96300000	1.03600000	-3.12500000
H	7.89500000	-0.01100000	-2.27100000
H	8.88600000	4.00000000	-1.10100000
H	5.65900000	0.24500000	-1.25700000
H	6.64600000	4.24600000	-0.05400000
H	4.36500000	1.55700000	0.12000000
H	-6.81000000	-0.60600000	-3.40000000
H	-6.98200000	-2.12500000	-4.33200000
H	-6.79300000	-3.00000000	-1.71700000
H	-4.66700000	-1.50000000	-4.10800000
H	-3.23200000	-1.72700000	-3.11000000
H	-4.79100000	-3.31400000	-0.47100000
H	-3.28000000	-2.93200000	-1.29900000
H	0.63300000	-4.69200000	2.11900000
H	1.54500000	-4.84500000	0.58800000
H	0.04400000	-3.91100000	-2.53100000
H	1.40900000	-4.35000000	-3.32900000
H	2.20100000	-2.19300000	-3.50800000
H	1.06900000	-0.26000000	-2.52900000
H	0.06100000	-1.05200000	-3.78400000
H	0.12800000	-0.72900000	3.57700000
H	-1.99400000	-1.94200000	3.88400000
H	-1.32500000	-3.37700000	3.05100000
H	1.23300000	-1.87300000	5.29300000
H	-0.49000000	2.03300000	3.83900000
H	0.66500000	1.32800000	3.02800000
H	2.55200000	0.11100000	3.46400000
H	2.71200000	0.80400000	2.12200000
H	-4.07100000	2.80400000	-2.26400000
H	-8.14900000	-1.71300000	-3.06400000
H	10.33000000	2.73000000	-2.76500000
H	0.21200000	-5.96900000	0.94400000
H	-5.73500000	7.15100000	0.04400000
H	-1.44900000	0.48100000	2.34400000
C	1.73600000	4.86800000	-0.00700000
C	1.70600000	3.51100000	0.31000000
C	0.56100000	2.76300000	-0.01800000
C	-0.54300000	3.36000000	-0.65800000
C	-0.49600000	4.71800000	-0.95900000
C	0.64000000	5.47000000	-0.63400000
H	2.62100000	5.44800000	0.23100000
H	2.56100000	3.04700000	0.78600000
H	-1.40100000	2.74900000	-0.91600000
H	-1.34000000	5.19200000	-1.44900000
H	0.67400000	6.52800000	-0.87800000
C	0.49900000	1.37100000	0.27200000
N	0.42500000	0.23400000	0.47800000

Table S3. Cartesian coordinates for TS1 CoNH

C	0.61800000	-4.63500000	1.96400000
S	-0.48300000	-3.82200000	0.78800000
N	0.86700000	-4.28600000	-2.54200000
C	1.52200000	-2.98000000	-2.50400000
C	0.53900000	-1.86200000	-2.81200000
S	-0.75300000	-1.64100000	-1.57400000
C	2.45800000	-2.77600000	-1.30000000
O	3.54900000	-3.36000000	-1.34800000
O	-1.43900000	-0.35000000	-1.92400000
O	-1.70700000	-2.82900000	-1.73300000

N	2.10000000	-1.95200000	-0.27900000
C	3.13900000	-1.75900000	0.75000000
C	2.51100000	-1.44100000	2.09000000
O	3.16900000	-1.01500000	3.05800000
C	4.06200000	-0.62300000	0.33700000
O	3.30600000	0.57900000	0.11700000
N	1.19400000	-1.67400000	2.12900000
C	0.45900000	-1.19500000	3.29100000
C	-0.88100000	-1.90700000	3.30500000
S	-1.54600000	-2.05600000	1.60600000
C	1.17700000	-1.51500000	4.61700000
O	1.32000000	-2.66100000	4.98500000
O	-1.90300000	0.49200000	1.72700000
C	-5.99800000	6.50200000	-0.48100000
C	-5.52900000	5.05800000	-0.36200000
C	-4.30900000	4.80500000	-1.24000000
C	-5.24300000	4.69800000	1.09200000
C	-4.09600000	1.04600000	-2.67600000
N	-3.80900000	0.84900000	-1.27000000
C	-4.66500000	1.13600000	-0.29500000
N	-5.88600000	1.64100000	-0.62800000
N	-4.38300000	0.84200000	0.97300000
C	9.82700000	1.56400000	-2.82600000
C	8.47400000	1.92000000	-2.28600000
C	7.40800000	1.01900000	-2.34000000
C	8.23400000	3.16400000	-1.68700000
C	6.16000000	1.33300000	-1.80800000
C	6.99800000	3.49100000	-1.14200000
C	5.95100000	2.56700000	-1.18800000
O	4.75300000	2.91100000	-0.63100000
C	-7.09900000	-2.57500000	-3.02100000
N	-6.28400000	-2.53100000	-1.81800000
C	-4.94100000	-2.61500000	-1.81100000
N	-4.24600000	-2.66000000	-2.94900000
N	-4.27200000	-2.62100000	-0.65700000
Co	0.30500000	-1.69300000	0.41900000
O	-0.40700000	1.87500000	3.87100000
O	2.20000000	1.65000000	2.87500000
H	4.77700000	-0.40400000	1.13500000
H	4.60800000	-0.90700000	-0.57000000
H	2.48900000	0.30600000	-0.33500000
H	3.74900000	-2.66800000	0.84900000
H	-5.21400000	7.18900000	-0.14000000
H	-6.88800000	6.69200000	0.12700000
H	-6.35200000	4.41500000	-0.71900000
H	-3.49300000	5.48800000	-0.97700000
H	-3.91200000	3.78900000	-1.11600000
H	-4.53200000	4.96400000	-2.30100000
H	-4.49600000	5.37700000	1.51800000
H	-4.83500000	3.68300000	1.19000000
H	-6.13800000	4.77700000	1.71800000
H	-5.00700000	0.51400000	-2.97700000
H	-3.25400000	0.63900000	-3.23400000
H	-2.90300000	0.41500000	-1.06600000
H	-5.01500000	1.20700000	1.67100000
H	-3.36200000	0.68700000	1.29600000
H	-6.39300000	2.11500000	0.10400000
H	-5.96000000	2.10800000	-1.52000000
H	10.55200000	1.39700000	-2.02100000
H	9.79200000	0.65000000	-3.42500000

H	7.55400000	0.04800000	-2.81000000
H	9.04100000	3.89300000	-1.63500000
H	5.33900000	0.62400000	-1.88200000
H	6.83100000	4.45400000	-0.66800000
H	4.22800000	2.10000000	-0.50300000
H	-6.74800000	-1.83100000	-3.74200000
H	-7.09600000	-3.56500000	-3.49000000
H	-6.74700000	-2.70500000	-0.94000000
H	-4.68500000	-2.95700000	-3.80500000
H	-3.22400000	-2.72200000	-2.86200000
H	-4.70100000	-2.31200000	0.19900000
H	-3.24900000	-2.68900000	-0.72000000
H	0.89800000	-3.99000000	2.80200000
H	1.52200000	-4.85700000	1.38800000
H	0.15000000	-4.36500000	-1.82600000
H	1.56300000	-5.00500000	-2.37600000
H	2.22200000	-2.97500000	-3.35000000
H	1.04200000	-0.89100000	-2.85200000
H	0.02200000	-2.04400000	-3.76000000
H	0.30000000	-0.11600000	3.21300000
H	-1.65000000	-1.36900000	3.86500000
H	-0.79500000	-2.92100000	3.71100000
H	1.44900000	-0.64700000	5.26800000
H	-0.53500000	2.69800000	4.35100000
H	0.50300000	1.92800000	3.50700000
H	2.64600000	0.87700000	3.26000000
H	2.48700000	1.60600000	1.95000000
H	-4.19600000	2.10900000	-2.92900000
H	-8.12300000	-2.32300000	-2.75500000
H	10.23600000	2.36000000	-3.45700000
H	0.14900000	-5.55600000	2.31300000
H	-6.23700000	6.76500000	-1.51500000
H	-1.66300000	0.90300000	2.57900000
C	-1.38900000	4.71700000	0.98300000
C	-1.45800000	3.32900000	1.04100000
C	-0.35100000	2.56300000	0.66100000
C	0.81700000	3.20900000	0.21300000
C	0.87600000	4.59400000	0.16000000
C	-0.22800000	5.35300000	0.54700000
H	-2.24800000	5.30800000	1.28800000
H	-2.35400000	2.83000000	1.39800000
H	1.67800000	2.61500000	-0.08100000
H	1.79000000	5.07400000	-0.17600000
H	-0.18000000	6.43700000	0.51500000
C	-0.29200000	1.12000000	0.69300000
N	0.37900000	0.19500000	0.34600000

Table S4. Cartesian coordinates for INT1 CoNH

C	0.44100000	-4.89800000	1.50700000
S	-0.63100000	-3.97300000	0.36500000
N	0.96700000	-3.83500000	-2.92400000
C	1.60500000	-2.54800000	-2.63600000
C	0.63400000	-1.38400000	-2.84900000
S	-0.74100000	-1.33200000	-1.67100000
C	2.45100000	-2.51700000	-1.33600000
O	3.58700000	-3.01200000	-1.42300000
O	-1.39700000	0.01600000	-1.90400000
O	-1.71200000	-2.46800000	-2.07300000
N	1.98300000	-1.95800000	-0.19400000

C	2.96500000	-1.83100000	0.90200000
C	2.24500000	-1.68100000	2.23400000
O	2.84500000	-1.41900000	3.29700000
C	3.89700000	-0.64000000	0.66400000
O	3.16400000	0.59500000	0.54900000
N	0.92400000	-1.86300000	2.14400000
C	0.11400000	-1.66200000	3.34400000
C	-1.18600000	-2.45500000	3.16500000
S	-1.76300000	-2.44300000	1.40800000
C	0.80100000	-2.17200000	4.64200000
O	0.93600000	-3.36000000	4.85400000
O	-1.33200000	0.75300000	1.78600000
C	-5.91900000	6.57600000	0.20400000
C	-5.33500000	5.15500000	0.19400000
C	-4.22400000	5.01300000	-0.85700000
C	-4.82600000	4.75800000	1.58900000
C	-4.10600000	1.34300000	-2.54200000
N	-3.69300000	1.06800000	-1.17200000
C	-4.44400000	1.26400000	-0.10300000
N	-5.72600000	1.70600000	-0.23900000
N	-3.98700000	0.95500000	1.12100000
C	9.82500000	1.57300000	-2.52300000
C	8.40000000	1.86400000	-2.09800000
C	7.41100000	0.87100000	-2.12300000
C	8.02500000	3.13600000	-1.63400000
C	6.10500000	1.12600000	-1.69100000
C	6.73000000	3.40200000	-1.18700000
C	5.76100000	2.39100000	-1.19800000
O	4.50000000	2.69500000	-0.74700000
C	-7.18000000	-2.15500000	-3.29400000
N	-6.30200000	-2.47700000	-2.16400000
C	-4.95400000	-2.45200000	-2.18600000
N	-4.27900000	-2.11400000	-3.29400000
N	-4.25100000	-2.72700000	-1.08400000
Co	0.19200000	-1.65900000	0.39300000
O	-0.21000000	1.71800000	3.92900000
O	2.30800000	1.34900000	3.24500000
H	4.58400000	-0.53600000	1.50700000
H	4.47000000	-0.82500000	-0.24700000
H	2.24700000	0.41000000	0.22200000
H	3.59300000	-2.72900000	0.95700000
H	-5.14600000	7.30700000	0.47100000
H	-6.73200000	6.66900000	0.93200000
H	-6.15500000	4.46900000	-0.07900000
H	-3.39900000	5.70300000	-0.64600000
H	-3.79300000	4.00400000	-0.86200000
H	-4.59400000	5.24000000	-1.86300000
H	-4.04100000	5.44800000	1.92200000
H	-4.38500000	3.75300000	1.58800000
H	-5.62700000	4.78900000	2.33600000
H	-4.99100000	0.75800000	-2.82000000
H	-3.27700000	1.04800000	-3.18300000
H	-2.74600000	0.64500000	-1.10700000
H	-4.54300000	1.21700000	1.92200000
H	-2.96800000	0.85100000	1.29700000
H	-6.16600000	2.12400000	0.56800000
H	-5.98400000	2.14400000	-1.11100000
H	10.46300000	1.37000000	-1.65300000
H	9.87600000	0.69900000	-3.17900000
H	7.65700000	-0.12200000	-2.49100000

H	8.76400000	3.93300000	-1.61000000
H	5.35300000	0.34500000	-1.75000000
H	6.45600000	4.38400000	-0.81300000
H	4.06500000	1.90300000	-0.36900000
H	-6.91000000	-1.18400000	-3.71900000
H	-7.13900000	-2.92300000	-4.07500000
H	-6.73200000	-2.90100000	-1.35400000
H	-4.71200000	-2.17300000	-4.20300000
H	-3.25300000	-2.12700000	-3.22800000
H	-4.68400000	-2.81400000	-0.17900000
H	-3.22000000	-2.71600000	-1.17300000
H	0.94600000	-4.22500000	2.20100000
H	1.17800000	-5.34500000	0.83200000
H	0.17000000	-4.02400000	-2.32400000
H	1.64500000	-4.58700000	-2.83900000
H	2.37100000	-2.41200000	-3.40800000
H	1.13200000	-0.42200000	-2.70600000
H	0.17500000	-1.43200000	-3.84100000
H	-0.11100000	-0.60300000	3.47600000
H	-2.02200000	-2.03300000	3.73000000
H	-1.06000000	-3.49900000	3.45900000
H	1.04400000	-1.40500000	5.40900000
H	-0.31700000	2.61100000	4.28000000
H	0.75900000	1.63000000	3.67500000
H	2.62600000	0.47800000	3.55100000
H	2.59300000	1.37300000	2.31100000
H	-4.30300000	2.41100000	-2.69500000
H	-8.20200000	-2.09300000	-2.92300000
H	10.26400000	2.42200000	-3.05800000
H	-0.11700000	-5.67700000	2.02800000
H	-6.31200000	6.85100000	-0.78000000
H	-1.08200000	1.19200000	2.65800000
C	-0.82900000	4.84100000	0.92300000
C	-1.01300000	3.47700000	1.15900000
C	-0.10600000	2.53300000	0.64800000
C	0.97800000	2.99100000	-0.12200000
C	1.17000000	4.35300000	-0.34500000
C	0.26700000	5.28500000	0.17800000
H	-1.53800000	5.55700000	1.32900000
H	-1.86600000	3.15200000	1.74500000
H	1.67700000	2.28400000	-0.54500000
H	2.03300000	4.67600000	-0.91800000
H	0.41800000	6.34700000	0.00900000
C	-0.28000000	1.07100000	0.90100000
N	0.43800000	0.19500000	0.34800000

Table S5. Cartesian coordinates for TS2 CoNH

C	0.47000000	-5.00700000	1.57800000
S	-0.63800000	-4.12700000	0.42700000
N	0.99800000	-3.98200000	-2.82200000
C	1.59500000	-2.66400000	-2.59000000
C	0.58100000	-1.54300000	-2.82700000
S	-0.79700000	-1.49900000	-1.64700000
C	2.45500000	-2.55000000	-1.30400000
O	3.63400000	-2.93600000	-1.40300000
O	-1.49500000	-0.18400000	-1.91900000
O	-1.72800000	-2.67900000	-2.02100000
N	1.94800000	-2.04700000	-0.15500000
C	2.92100000	-1.83500000	0.93300000

C	2.19300000	-1.69600000	2.26200000
O	2.78600000	-1.36700000	3.32000000
C	3.79100000	-0.60000000	0.67100000
O	3.00300000	0.60800000	0.57400000
N	0.88600000	-1.93200000	2.18200000
C	0.06500000	-1.76300000	3.38200000
C	-1.21200000	-2.59200000	3.19800000
S	-1.78100000	-2.57800000	1.44100000
C	0.76100000	-2.24900000	4.68300000
O	0.97100000	-3.42800000	4.87800000
O	-1.30700000	0.77700000	1.77200000
C	-6.21500000	6.27100000	0.19900000
C	-4.75000000	5.81800000	0.29600000
C	-3.78800000	6.96200000	-0.06000000
C	-4.43600000	5.26100000	1.69500000
C	-4.23700000	1.08300000	-2.51900000
N	-3.69500000	1.12400000	-1.16500000
C	-4.26900000	1.75500000	-0.14700000
N	-5.45300000	2.39900000	-0.34500000
N	-3.75100000	1.70100000	1.08300000
C	9.68200000	1.72100000	-2.42900000
C	8.23400000	1.96900000	-2.05500000
C	7.27900000	0.94300000	-2.09900000
C	7.80200000	3.23400000	-1.62400000
C	5.95000000	1.16200000	-1.71900000
C	6.48400000	3.46400000	-1.22800000
C	5.54900000	2.42200000	-1.26000000
O	4.26000000	2.69700000	-0.86900000
C	-7.20200000	-2.50600000	-3.27300000
N	-6.32100000	-2.76300000	-2.13000000
C	-4.97300000	-2.72500000	-2.15200000
N	-4.29800000	-2.44000000	-3.27500000
N	-4.26800000	-2.94200000	-1.03800000
Co	0.15100000	-1.77000000	0.42000000
O	-0.01300000	1.57400000	3.67600000
O	2.32100000	1.26000000	3.11700000
H	4.49800000	-0.46400000	1.49300000
H	4.34500000	-0.75800000	-0.25500000
H	2.07500000	0.38800000	0.24700000
H	3.60200000	-2.69300000	1.00800000
H	-6.42000000	7.07600000	0.91400000
H	-6.91200000	5.45200000	0.42300000
H	-4.58600000	5.01400000	-0.43700000
H	-3.89400000	7.79100000	0.65100000
H	-2.74700000	6.62600000	-0.03800000
H	-3.99400000	7.35400000	-1.06200000
H	-4.51000000	6.05000000	2.45200000
H	-3.42300000	4.84600000	1.73900000
H	-5.15100000	4.47900000	1.99600000
H	-5.24700000	0.65900000	-2.52900000
H	-3.57500000	0.44400000	-3.10000000
H	-2.77100000	0.67300000	-1.06900000
H	-4.11300000	2.33900000	1.77600000
H	-2.80100000	1.29500000	1.28800000
H	-5.79100000	3.01400000	0.38100000
H	-5.69900000	2.68700000	-1.28100000
H	9.78400000	0.83700000	-3.06700000
H	10.10800000	2.57500000	-2.96500000
H	7.57000000	-0.04600000	-2.44300000
H	8.51400000	4.05600000	-1.58700000

H	5.22600000	0.35700000	-1.79600000
H	6.16500000	4.44200000	-0.88100000
H	3.84400000	1.91600000	-0.45100000
H	-6.94400000	-1.55400000	-3.74500000
H	-7.15300000	-3.31100000	-4.01500000
H	-6.74600000	-3.14400000	-1.29700000
H	-4.73100000	-2.53600000	-4.18000000
H	-3.27200000	-2.43500000	-3.20600000
H	-4.70200000	-2.97500000	-0.12900000
H	-3.23800000	-2.90900000	-1.12500000
H	0.98600000	-4.31000000	2.23800000
H	1.19200000	-5.47100000	0.89900000
H	0.21400000	-4.17200000	-2.20700000
H	1.70000000	-4.70800000	-2.71900000
H	2.34600000	-2.52900000	-3.37700000
H	1.04700000	-0.56100000	-2.71500000
H	0.11900000	-1.63300000	-3.81500000
H	-0.20500000	-0.71400000	3.50500000
H	-2.05600000	-2.19400000	3.76800000
H	-1.05800000	-3.63500000	3.48400000
H	0.95000000	-1.48400000	5.47000000
H	-0.16500000	2.50100000	3.90200000
H	1.11200000	1.44100000	3.37500000
H	2.59600000	0.36400000	3.44300000
H	2.55500000	1.22200000	2.14700000
H	-4.25100000	2.07700000	-2.98400000
H	-8.22500000	-2.43700000	-2.90300000
H	10.29800000	1.55500000	-1.53600000
H	-0.06800000	-5.77200000	2.13800000
H	-6.45000000	6.64800000	-0.80200000
H	-0.70900000	1.23700000	2.83200000
C	-0.48800000	4.80400000	0.82000000
C	-0.75700000	3.46900000	1.13300000
C	-0.10100000	2.42300000	0.46000000
C	0.80800000	2.75500000	-0.55900000
C	1.08800000	4.08600000	-0.86500000
C	0.44200000	5.11700000	-0.17400000
H	-0.99200000	5.59700000	1.36400000
H	-1.47900000	3.23900000	1.90900000
H	1.30100000	1.96300000	-1.10800000
H	1.82800000	4.31000000	-1.62700000
H	0.66700000	6.15400000	-0.40600000
C	-0.34200000	0.98300000	0.84000000
N	0.39500000	0.08500000	0.31200000

Table S6. Cartesian coordinates for INT2 CoNH

C	0.77400000	-4.64500000	2.05200000
S	-0.25200000	-3.92100000	0.74300000
N	1.21100000	-4.11000000	-2.52800000
C	1.78000000	-2.76400000	-2.44000000
C	0.74400000	-1.70200000	-2.80300000
S	-0.57600000	-1.48900000	-1.57200000
C	2.62500000	-2.49600000	-1.16400000
O	3.72600000	-3.06100000	-1.12200000
O	-1.22900000	-0.17400000	-1.93700000
O	-1.54600000	-2.67400000	-1.77300000
N	2.18700000	-1.62300000	-0.20700000
C	3.17300000	-1.32100000	0.85300000
C	2.45100000	-1.04200000	2.16500000

O	2.98300000	-0.41900000	3.10600000
C	4.05100000	-0.14100000	0.42200000
O	3.24400000	0.90400000	-0.16300000
N	1.21200000	-1.55700000	2.20600000
C	0.36900000	-1.18400000	3.34700000
C	-0.90300000	-2.03500000	3.29300000
S	-1.45400000	-2.29400000	1.54300000
C	1.07800000	-1.44200000	4.70100000
O	1.34500000	-2.57100000	5.05700000
O	-1.97600000	0.25600000	1.32000000
C	-5.92700000	6.28300000	-1.02400000
C	-5.43800000	4.85200000	-0.76000000
C	-4.12900000	4.56200000	-1.51100000
C	-5.27400000	4.59800000	0.74600000
C	-3.95800000	0.73700000	-2.91500000
N	-3.76100000	0.69500000	-1.47100000
C	-4.68800000	0.99700000	-0.57200000
N	-5.92100000	1.40200000	-0.99000000
N	-4.45500000	0.80500000	0.73300000
C	9.96000000	1.38100000	-2.99700000
C	8.64400000	1.79900000	-2.37600000
C	7.46900000	1.06200000	-2.56700000
C	8.56200000	2.94900000	-1.57200000
C	6.25800000	1.44700000	-1.97800000
C	7.36700000	3.34200000	-0.97200000
C	6.20300000	2.58600000	-1.16600000
O	5.05700000	3.01500000	-0.55300000
C	-6.92500000	-2.92600000	-3.08800000
N	-6.11700000	-2.67900000	-1.89300000
C	-4.77000000	-2.69600000	-1.85900000
N	-4.05000000	-2.92600000	-2.96300000
N	-4.12100000	-2.46100000	-0.71300000
Co	0.38500000	-1.59000000	0.48500000
O	-0.74300000	1.56500000	4.74200000
O	1.57200000	1.97000000	3.25900000
H	4.55300000	0.29500000	1.28700000
H	4.79900000	-0.48300000	-0.30100000
H	2.73000000	0.46700000	-0.86200000
H	3.83200000	-2.18200000	1.02300000
H	-5.19200000	7.01300000	-0.66400000
H	-6.87500000	6.48300000	-0.51200000
H	-6.21300000	4.16800000	-1.14300000
H	-3.33500000	5.23900000	-1.17700000
H	-3.76600000	3.54400000	-1.32600000
H	-4.25000000	4.69900000	-2.59200000
H	-4.54100000	5.29100000	1.17400000
H	-4.89600000	3.58900000	0.94800000
H	-6.21900000	4.73700000	1.28400000
H	-4.80700000	0.11200000	-3.21600000
H	-3.04800000	0.34800000	-3.36700000
H	-2.82100000	0.39500000	-1.18100000
H	-5.07900000	1.24700000	1.39500000
H	-3.44900000	0.65500000	1.04700000
H	-6.52200000	1.83900000	-0.30500000
H	-5.99500000	1.81100000	-1.91100000
H	10.68800000	1.09200000	-2.23000000
H	9.83100000	0.52700000	-3.67000000
H	7.49100000	0.17000000	-3.18800000
H	9.45600000	3.54800000	-1.40800000
H	5.35700000	0.86600000	-2.15800000

H	7.31700000	4.22700000	-0.34600000
H	4.37400000	2.31700000	-0.56300000
H	-6.61200000	-2.26400000	-3.90100000
H	-6.86200000	-3.97000000	-3.41700000
H	-6.58900000	-2.69400000	-1.00000000
H	-4.45700000	-3.37200000	-3.77100000
H	-3.02500000	-2.87900000	-2.87000000
H	-4.56000000	-1.96900000	0.05200000
H	-3.09600000	-2.48800000	-0.76200000
H	1.25600000	-3.86500000	2.64300000
H	1.53400000	-5.21600000	1.50900000
H	0.43200000	-4.24500000	-1.89100000
H	1.93000000	-4.80100000	-2.33100000
H	2.53500000	-2.70300000	-3.23300000
H	1.19900000	-0.71100000	-2.87800000
H	0.23800000	-1.95100000	-3.74100000
H	0.09500000	-0.13100000	3.28600000
H	-1.74800000	-1.53500000	3.77200000
H	-0.76000000	-3.02200000	3.74000000
H	1.27100000	-0.56100000	5.34800000
H	-0.72100000	1.91800000	5.63900000
H	0.08500000	1.86700000	4.31000000
H	2.20500000	1.22100000	3.29700000
H	2.08500000	2.76000000	3.04700000
H	-4.11200000	1.76300000	-3.27000000
H	-7.96500000	-2.70100000	-2.84900000
H	10.40600000	2.19800000	-3.57600000
H	0.18100000	-5.31600000	2.67400000
H	-6.07700000	6.46200000	-2.09500000
H	1.01300000	0.87800000	0.51400000
C	-1.65700000	4.48100000	1.91000000
C	-1.61300000	3.08700000	1.94300000
C	-0.98200000	2.38800000	0.90400000
C	-0.41900000	3.08600000	-0.17400000
C	-0.47900000	4.48100000	-0.20900000
C	-1.09300000	5.17900000	0.83500000
H	-2.12000000	5.02500000	2.72800000
H	-2.00400000	2.53800000	2.79400000
H	0.04300000	2.53500000	-0.98700000
H	-0.04500000	5.02000000	-1.04600000
H	-1.12700000	6.26500000	0.81500000
C	-0.92300000	0.89600000	0.95700000
N	0.19100000	0.28700000	0.61500000

Table S7. Cartesian coordinates for TS3 CoNH

C	1.02700000	-5.07500000	0.49500000
S	-0.12400000	-3.91800000	-0.30300000
N	1.19800000	-3.36300000	-3.40600000
C	1.84000000	-2.19900000	-2.80100000
C	0.95500000	-0.96800000	-2.88700000
S	-0.43900000	-0.95400000	-1.72100000
C	2.47800000	-2.45400000	-1.41600000
O	3.27000000	-3.40200000	-1.35500000
O	-0.95400000	0.44900000	-1.83900000
O	-1.43700000	-2.00900000	-2.24500000
N	2.21900000	-1.57700000	-0.37400000
C	3.19800000	-1.70100000	0.73200000
C	2.49000000	-1.83700000	2.06500000
O	3.05200000	-1.67300000	3.15600000

C	4.17700000	-0.52800000	0.67900000
O	3.47000000	0.69800000	0.44400000
N	1.21800000	-2.26200000	1.92500000
C	0.46500000	-2.51800000	3.14700000
C	-0.69700000	-3.44500000	2.78500000
S	-1.42300000	-3.00100000	1.13800000
C	1.35000000	-3.19400000	4.22300000
O	1.66100000	-4.36400000	4.13800000
O	-2.04300000	1.05100000	0.62500000
C	-5.70000000	6.16300000	1.98300000
C	-5.08800000	4.92900000	1.30200000
C	-4.49500000	5.29400000	-0.06700000
C	-4.03200000	4.25600000	2.19200000
C	-3.79700000	1.77200000	-1.92800000
N	-4.35500000	1.51800000	-0.60600000
C	-5.59100000	1.78500000	-0.22600000
N	-6.42000000	2.48400000	-1.04200000
N	-6.05000000	1.29300000	0.93700000
C	10.11800000	2.47000000	-2.07600000
C	8.75300000	2.61100000	-1.43600000
C	7.71200000	1.72000000	-1.72600000
C	8.49200000	3.62900000	-0.50500000
C	6.46100000	1.82800000	-1.10800000
C	7.25300000	3.74800000	0.12600000
C	6.22900000	2.84000000	-0.16700000
O	5.02900000	2.99200000	0.47500000
C	-6.78100000	-1.56200000	-3.42000000
N	-5.97200000	-1.51000000	-2.20300000
C	-4.62900000	-1.65800000	-2.19100000
N	-3.94700000	-1.73100000	-3.34300000
N	-3.96000000	-1.71200000	-1.04100000
Co	0.35300000	-1.71500000	0.30400000
O	-4.43000000	-0.84500000	1.67400000
O	-2.10700000	-0.90000000	2.67400000
H	4.69500000	-0.42800000	1.63500000
H	4.91400000	-0.69500000	-0.11600000
H	2.82600000	0.49700000	-0.25700000
H	3.77500000	-2.62700000	0.62400000
H	-4.92900000	6.91900000	2.17400000
H	-6.15300000	5.90300000	2.94600000
H	-5.90400000	4.20700000	1.13800000
H	-3.69200000	6.03100000	0.04600000
H	-4.05900000	4.41600000	-0.55700000
H	-5.24900000	5.73000000	-0.73300000
H	-3.19900000	4.94200000	2.38500000
H	-3.61200000	3.36000000	1.71800000
H	-4.45400000	3.96400000	3.16000000
H	-4.44400000	1.34900000	-2.70300000
H	-2.81800000	1.29400000	-1.96700000
H	-3.66400000	1.21000000	0.08700000
H	-6.80000000	1.75700000	1.42600000
H	-5.47000000	0.59600000	1.43200000
H	-7.35900000	2.68400000	-0.73400000
H	-6.01500000	3.15400000	-1.68100000
H	10.07000000	1.87600000	-2.99400000
H	10.54300000	3.44700000	-2.33000000
H	7.87400000	0.92500000	-2.45000000
H	9.27600000	4.34300000	-0.26200000
H	5.66400000	1.13600000	-1.36400000
H	7.06600000	4.53300000	0.85100000

H	4.47700000	2.19200000	0.36400000
H	-6.46400000	-0.78800000	-4.12500000
H	-6.73400000	-2.54500000	-3.90400000
H	-6.44200000	-1.63800000	-1.31800000
H	-4.40800000	-1.91600000	-4.21900000
H	-2.92300000	-1.80000000	-3.27900000
H	-4.32600000	-1.41200000	-0.13200000
H	-2.94900000	-1.83100000	-1.13500000
H	1.40400000	-4.68300000	1.43900000
H	1.85900000	-5.12900000	-0.21800000
H	0.27400000	-3.53600000	-3.02300000
H	1.78200000	-4.18700000	-3.29400000
H	2.72200000	-1.97200000	-3.41900000
H	1.50700000	-0.05900000	-2.64400000
H	0.49600000	-0.87100000	-3.87500000
H	0.08100000	-1.58600000	3.57700000
H	-1.54100000	-3.35400000	3.46800000
H	-0.37500000	-4.48700000	2.74000000
H	1.61600000	-2.56000000	5.08900000
H	-4.85900000	-1.51300000	2.22500000
H	-3.46000000	-0.77700000	2.06500000
H	-2.10100000	-0.66600000	3.61100000
H	-1.06300000	-0.44400000	2.07700000
H	-3.66700000	2.84500000	-2.11700000
H	-7.81900000	-1.36200000	-3.15000000
H	10.82400000	1.97300000	-1.39900000
H	0.54400000	-6.04600000	0.61100000
H	-6.47200000	6.62500000	1.35800000
H	0.69500000	0.26000000	1.81500000
C	-0.19200000	4.70300000	-0.22000000
C	-0.81200000	3.47900000	0.00700000
C	-0.11400000	2.43200000	0.63700000
C	1.21200000	2.64400000	1.04600000
C	1.83800000	3.87100000	0.81400000
C	1.13700000	4.90000000	0.18000000
H	-0.73600000	5.50200000	-0.71600000
H	-1.83400000	3.31000000	-0.30700000
H	1.78900000	1.86300000	1.52900000
H	2.87300000	4.01100000	1.10800000
H	1.62600000	5.85100000	-0.00700000
C	-0.83200000	1.14800000	0.83300000
N	-0.16800000	0.00600000	1.33800000

Table S8. Cartesian coordinates for EP CoNH

C	0.91700000	-5.09900000	-0.14700000
S	-0.22800000	-3.87600000	-0.87400000
N	1.38100000	-2.13600000	-3.72700000
C	1.89500000	-1.20900000	-2.71800000
C	0.88500000	-0.10600000	-2.38900000
S	-0.54400000	-0.73100000	-1.46400000
C	2.53900000	-1.89900000	-1.49200000
O	3.57800000	-2.52900000	-1.68100000
O	-1.22600000	0.48900000	-0.87200000
O	-1.40000000	-1.43400000	-2.52200000
N	1.99700000	-1.69500000	-0.23100000
C	2.90800000	-2.06000000	0.86900000
C	2.10400000	-2.30700000	2.13200000
O	2.61800000	-2.38500000	3.25400000
C	3.92700000	-0.92100000	1.06000000

O	3.23400000	0.34000000	0.88200000
N	0.78000000	-2.45900000	1.88900000
C	-0.07800000	-2.86800000	3.00800000
C	-1.33500000	-3.49900000	2.40500000
S	-1.93400000	-2.49500000	0.99000000
C	0.64000000	-3.89300000	3.91500000
O	0.75500000	-5.05500000	3.58400000
O	-1.31200000	2.11100000	2.18400000
C	-6.02600000	5.72300000	2.76300000
C	-5.23800000	4.83800000	1.78300000
C	-4.23900000	5.67000000	0.96400000
C	-4.52300000	3.69300000	2.51700000
C	-3.69800000	2.13400000	-1.69900000
N	-3.80500000	1.22100000	-0.57000000
C	-4.90500000	1.06100000	0.15000000
N	-6.00600000	1.80800000	-0.13000000
N	-4.97700000	0.11000000	1.09100000
C	10.15300000	2.70200000	-0.30400000
C	8.66000000	2.51700000	-0.45500000
C	7.89200000	1.92600000	0.55600000
C	7.99700000	2.90900000	-1.63000000
C	6.51700000	1.72500000	0.40700000
C	6.62600000	2.71300000	-1.79800000
C	5.88100000	2.11500000	-0.77700000
O	4.53200000	1.94400000	-0.97400000
C	-6.48600000	-0.82400000	-4.09200000
N	-5.78100000	-1.77000000	-3.23400000
C	-4.45400000	-1.70400000	-3.01000000
N	-3.72400000	-0.75200000	-3.60400000
N	-3.86000000	-2.55200000	-2.16600000
Co	0.16200000	-1.97700000	0.21500000
O	-1.60100000	0.10800000	3.71200000
O	-2.73800000	-1.27100000	1.85200000
H	4.34200000	-0.94700000	2.06900000
H	4.72400000	-0.99900000	0.31500000
H	2.52000000	0.09000000	0.25400000
H	3.45700000	-2.98300000	0.64700000
H	-5.34700000	6.19600000	3.48200000
H	-6.75800000	5.13700000	3.33200000
H	-5.96700000	4.39800000	1.08100000
H	-3.48200000	6.11700000	1.61900000
H	-3.70600000	5.05300000	0.23000000
H	-4.73900000	6.48300000	0.42600000
H	-3.82700000	4.09100000	3.26400000
H	-3.92500000	3.07100000	1.84100000
H	-5.23800000	3.04900000	3.04600000
H	-4.44800000	1.90500000	-2.46400000
H	-2.70000000	2.00400000	-2.11400000
H	-2.93300000	0.71800000	-0.34700000
H	-5.71500000	0.15000000	1.77900000
H	-4.13700000	-0.43900000	1.34000000
H	-6.78400000	1.77300000	0.51200000
H	-5.87500000	2.70500000	-0.57600000
H	10.48900000	3.64000000	-0.75800000
H	10.45100000	2.71300000	0.75000000
H	8.37200000	1.61800000	1.48200000
H	8.56500000	3.37600000	-2.43200000
H	5.94000000	1.27700000	1.21200000
H	6.12400000	3.01400000	-2.71200000
H	4.15900000	1.42100000	-0.23600000

H	-6.29700000	0.20200000	-3.76300000
H	-6.20500000	-0.93200000	-5.14700000
H	-6.27000000	-2.59900000	-2.93000000
H	-4.05400000	-0.31500000	-4.45000000
H	-2.71100000	-0.75600000	-3.39100000
H	-4.38900000	-3.17400000	-1.57400000
H	-2.84100000	-2.48400000	-2.06000000
H	0.79500000	-5.19000000	0.93400000
H	1.94900000	-4.82500000	-0.38000000
H	0.64400000	-2.72000000	-3.33800000
H	2.13600000	-2.75400000	-4.01500000
H	2.74600000	-0.68900000	-3.17500000
H	1.31900000	0.68100000	-1.76700000
H	0.47800000	0.33000000	-3.30600000
H	-0.33500000	-1.99100000	3.61500000
H	-2.16400000	-3.57600000	3.11300000
H	-1.10700000	-4.48800000	2.00400000
H	1.01500000	-3.50800000	4.87800000
H	-2.03700000	0.22100000	4.56600000
H	-1.57900000	1.00300000	3.22500000
H	-2.20600000	-0.81100000	2.58100000
H	0.51900000	0.67800000	3.05600000
H	-3.79900000	3.18000000	-1.38500000
H	-7.55600000	-1.01400000	-4.00700000
H	10.70500000	1.88800000	-0.79200000
H	0.69000000	-6.05500000	-0.62700000
H	-6.56500000	6.51900000	2.23800000
H	1.73200000	1.13800000	1.91700000
C	-0.53100000	4.35500000	-1.24300000
C	-0.80400000	3.59600000	-0.10700000
C	0.22800000	2.91200000	0.55200000
C	1.54000000	3.02100000	0.06900000
C	1.81300000	3.76700000	-1.07900000
C	0.77600000	4.43200000	-1.73900000
H	-1.33400000	4.89000000	-1.74300000
H	-1.81000000	3.51600000	0.28700000
H	2.36900000	2.54900000	0.58200000
H	2.83500000	3.81400000	-1.44000000
H	0.98700000	5.01900000	-2.62800000
C	-0.14100000	2.06900000	1.72600000
N	0.78400000	1.25100000	2.26500000

Table S9. Cartesian coordinates for E+P CoNH

C	-1.54100000	-5.04600000	0.67000000
S	-0.21000000	-3.86800000	1.11000000
N	-1.46500000	-1.76700000	3.85000000
C	-1.98000000	-0.88300000	2.80300000
C	-0.91500000	0.09900000	2.30800000
S	0.42400000	-0.69800000	1.38100000
C	-2.78500000	-1.60600000	1.69700000
O	-3.88000000	-2.06700000	2.02800000
O	1.18500000	0.44000000	0.72200000
O	1.26600000	-1.40100000	2.45300000
N	-2.31800000	-1.62400000	0.39900000
C	-3.33000000	-2.06400000	-0.58800000
C	-2.64600000	-2.66600000	-1.80600000
O	-3.25100000	-2.89800000	-2.86500000
C	-4.16100000	-0.83000000	-0.99000000
O	-3.25500000	0.29800000	-1.12300000

N	-1.33900000	-2.90700000	-1.61500000
C	-0.57000000	-3.40500000	-2.75300000
C	0.72900000	-3.98800000	-2.20800000
S	1.46500000	-2.80000000	-1.01700000
C	-1.34700000	-4.48200000	-3.53900000
O	-1.47300000	-5.61000000	-3.11400000
O	0.47100000	2.16600000	-1.58200000
C	6.46500000	4.32300000	-4.01200000
C	5.85300000	3.64900000	-2.77700000
C	5.94300000	4.56800000	-1.54900000
C	4.40200000	3.21700000	-3.03800000
C	3.97700000	1.69800000	1.00000000
N	3.75900000	0.75300000	-0.08600000
C	4.71600000	0.33600000	-0.90100000
N	5.98400000	0.80300000	-0.73900000
N	4.47400000	-0.59300000	-1.83400000
C	-9.76600000	3.79400000	0.06200000
C	-8.28900000	3.46800000	0.09100000
C	-7.73800000	2.52700000	-0.79100000
C	-7.42500000	4.07200000	1.01700000
C	-6.38200000	2.19200000	-0.74900000
C	-6.06800000	3.74700000	1.07600000
C	-5.54400000	2.80000000	0.19200000
O	-4.20200000	2.49900000	0.27400000
C	6.52200000	-1.23000000	3.68300000
N	5.63400000	-2.18100000	3.01900000
C	4.31400000	-1.96800000	2.84200000
N	3.74400000	-0.85000000	3.31100000
N	3.56500000	-2.84000000	2.16200000
Co	-0.51100000	-2.04300000	-0.15800000
O	-0.33500000	-0.31400000	-1.73000000
O	1.93900000	-1.60600000	-2.15700000
H	-4.64900000	-0.99300000	-1.95200000
H	-4.90300000	-0.61100000	-0.21500000
H	-2.57300000	0.07700000	-0.44800000
H	-4.00300000	-2.81800000	-0.16600000
H	5.91500000	5.23700000	-4.26500000
H	6.42900000	3.66200000	-4.88500000
H	6.44700000	2.74500000	-2.57100000
H	5.37800000	5.49200000	-1.72100000
H	5.52000000	4.09700000	-0.65300000
H	6.97900000	4.85100000	-1.33000000
H	3.78000000	4.08900000	-3.27400000
H	3.95400000	2.72600000	-2.16700000
H	4.33300000	2.52900000	-3.88700000
H	4.75700000	1.34200000	1.68200000
H	3.03600000	1.77500000	1.54200000
H	2.79000000	0.41300000	-0.16000000
H	5.15500000	-0.77000000	-2.55700000
H	3.51900000	-0.95400000	-1.97900000
H	6.66700000	0.58600000	-1.45000000
H	6.11000000	1.70600000	-0.30500000
H	-9.95600000	4.81900000	0.39900000
H	-10.17900000	3.68900000	-0.94600000
H	-8.37600000	2.04600000	-1.52800000
H	-7.82000000	4.81000000	1.71100000
H	-5.97600000	1.46900000	-1.45200000
H	-5.41100000	4.21600000	1.80200000
H	-4.00300000	1.72500000	-0.29500000
H	6.44200000	-0.24500000	3.21400000

H	6.31000000	-1.14700000	4.75600000
H	5.98900000	-3.10500000	2.82000000
H	4.17300000	-0.34300000	4.06900000
H	2.72700000	-0.75600000	3.14700000
H	3.97100000	-3.61300000	1.65800000
H	2.55600000	-2.65000000	2.09000000
H	-1.51100000	-5.31800000	-0.38600000
H	-2.51100000	-4.60800000	0.91300000
H	-0.80900000	-2.44000000	3.46000000
H	-2.24400000	-2.29300000	4.23800000
H	-2.74200000	-0.25200000	3.27700000
H	-1.33700000	0.83500000	1.62000000
H	-0.42600000	0.60000000	3.14900000
H	-0.35000000	-2.58300000	-3.45200000
H	1.48100000	-4.17000000	-2.98000000
H	0.53800000	-4.90800000	-1.65200000
H	-1.72500000	-4.15700000	-4.52300000
H	-1.16200000	-0.32200000	-2.23900000
H	-0.05400000	0.64600000	-1.63300000
H	1.14000000	-1.01900000	-2.25100000
H	-1.96900000	2.20600000	-1.12600000
H	4.22800000	2.69500000	0.62400000
H	7.54800000	-1.58000000	3.56600000
H	-10.33100000	3.12300000	0.72100000
H	-1.38300000	-5.93300000	1.28900000
H	7.51100000	4.60000000	-3.84000000
H	-1.93000000	3.30400000	0.18000000
C	2.64300000	5.12100000	0.53400000
C	1.85700000	4.30500000	-0.28000000
C	0.71900000	3.67600000	0.24500000
C	0.38700000	3.85200000	1.59600000
C	1.18700000	4.65200000	2.41400000
C	2.31200000	5.29300000	1.88300000
H	3.50800000	5.62800000	0.11700000
H	2.10600000	4.15100000	-1.32400000
H	-0.47500000	3.34400000	2.01900000
H	0.93000000	4.78000000	3.46100000
H	2.92200000	5.93000000	2.51700000
C	-0.08200000	2.80200000	-0.66000000
N	-1.41000000	2.70200000	-0.44300000

Table S10. Cartesian coordinates for E FeNH

C	3.86300000	-3.89300000	1.90500000
O	3.10800000	-2.94200000	2.18100000
N	5.12500000	-3.94100000	2.38200000
S	-4.01200000	-1.72100000	-1.98900000
N	-5.10800000	1.53700000	-2.27000000
C	-3.86400000	2.17000000	-1.84200000
C	-3.54800000	2.15900000	-0.32700000
S	-3.48800000	0.47600000	0.33800000
C	-2.70300000	1.68800000	-2.74300000
O	-2.49500000	2.33600000	-3.77600000
O	-2.70900000	0.59900000	1.68500000
O	-4.94500000	0.07100000	0.60400000
N	-1.92600000	0.62000000	-2.38400000
C	-0.83500000	0.34000000	-3.34800000
C	-0.31800000	-1.07200000	-3.14800000
O	0.64400000	-1.52500000	-3.82500000
C	0.33300000	1.34500000	-3.26900000

O	0.96200000	1.36900000	-1.97500000
N	-0.93800000	-1.75500000	-2.19700000
C	-0.48000000	-3.12500000	-1.94700000
C	-1.54700000	-3.85700000	-1.13000000
S	-2.22800000	-2.78000000	0.19100000
C	-0.22800000	-3.90300000	-3.25900000
O	-1.13900000	-4.30100000	-3.95500000
O	-0.92700000	-2.84600000	1.27300000
C	6.23100000	-4.12400000	-1.97000000
C	5.10300000	-3.43000000	-2.35300000
C	6.73600000	-3.44100000	-0.80300000
N	4.87800000	-2.36100000	-1.50400000
C	5.87300000	-2.33100000	-0.55400000
C	7.84300000	-3.64200000	0.04600000
C	6.10200000	-1.41700000	0.48300000
C	8.07000000	-2.74200000	1.08600000
C	7.21200000	-1.63600000	1.29500000
C	6.25600000	1.88100000	2.26400000
C	5.18100000	1.26500000	2.92000000
C	6.01400000	2.39700000	0.98000000
C	3.91500000	1.16600000	2.33200000
C	4.75500000	2.31400000	0.37800000
C	3.69700000	1.69700000	1.05800000
O	2.43300000	1.58200000	0.52800000
N	-3.38900000	-1.49200000	3.36100000
C	-2.62200000	-2.15200000	4.22100000
N	-3.16500000	-2.73600000	5.31400000
N	-1.30700000	-2.25100000	3.97800000
C	0.80700000	6.77200000	-2.94900000
C	-0.19900000	5.88000000	-3.35300000
C	1.39000000	6.56100000	-1.69200000
C	-0.60700000	4.81600000	-2.54800000
C	1.00200000	5.49500000	-0.87100000
C	0.00500000	4.62000000	-1.30700000
O	-0.40400000	3.53200000	-0.54700000
C	1.20400000	5.63700000	3.11200000
C	1.33300000	4.27900000	2.79700000
C	-0.10200000	6.15800000	3.18500000
C	0.20900000	3.45900000	2.58000000
C	-1.22500000	5.36500000	2.97500000
C	-1.08100000	3.99900000	2.68600000
O	-2.23600000	3.29000000	2.50800000
N	-8.02400000	-3.26600000	1.66500000
C	-7.37300000	-2.16800000	1.23200000
N	-7.60800000	-0.97500000	1.80900000
N	-6.46900000	-2.23500000	0.25800000
Fe	-2.18700000	-0.77300000	-1.10600000
H	5.80200000	-4.55800000	1.95400000
H	5.50600000	-3.07400000	2.74100000
H	1.10000000	1.03200000	-3.98000000
H	-0.02100000	2.34200000	-3.53200000
H	0.31800000	1.71900000	-1.33300000
H	-1.21600000	0.40800000	-4.37500000
H	4.44900000	-3.60400000	-3.19700000
H	4.11600000	-1.69200000	-1.59800000
H	8.52300000	-4.47300000	-0.12400000
H	8.93100000	-2.87700000	1.73500000
H	5.44600000	-0.56900000	0.64700000
H	7.42600000	-0.93400000	2.09600000
H	6.82400000	2.87300000	0.43300000

H	5.33000000	0.84400000	3.91200000
H	4.59600000	2.72400000	-0.61700000
H	3.09800000	0.66900000	2.84600000
H	2.39800000	1.92900000	-0.37800000
H	-2.95100000	-0.79600000	2.72900000
H	-2.60700000	-3.32900000	5.90800000
H	-4.04400000	-2.40800000	5.68500000
H	-0.68500000	-2.63100000	4.67600000
H	-1.01800000	-2.32100000	2.99300000
H	2.16500000	7.23700000	-1.33900000
H	-0.68200000	6.01800000	-4.31700000
H	1.46100000	5.35400000	0.10400000
H	-1.38900000	4.14500000	-2.88500000
H	-0.05600000	3.62300000	0.35800000
H	-0.24000000	7.21300000	3.41200000
H	2.32000000	3.83100000	2.71600000
H	-2.22500000	5.78100000	3.03400000
H	0.35400000	2.40400000	2.35200000
H	-2.10700000	2.34400000	2.31400000
H	-7.89800000	-4.11000000	1.12500000
H	-6.14300000	-3.11700000	-0.11300000
H	-5.87500000	-1.41500000	0.05600000
H	-8.43200000	-0.81500000	2.36700000
H	-7.08000000	-0.16700000	1.49200000
H	-5.90900000	1.95400000	-1.80500000
H	-5.10700000	0.53600000	-2.08900000
H	-3.94000000	3.23100000	-2.09900000
H	-2.56800000	2.60400000	-0.13200000
H	-4.31200000	2.68300000	0.25400000
H	0.46700000	-3.10600000	-1.39400000
H	-1.15900000	-4.75100000	-0.63600000
H	-2.39400000	-4.12500000	-1.76300000
H	0.82600000	-4.13200000	-3.49600000
C	7.63000000	1.98800000	2.90100000
H	8.39200000	1.50300000	2.28000000
H	7.93000000	3.03400000	3.02700000
H	7.64800000	1.51500000	3.88800000
C	2.39700000	6.53400000	3.39200000
H	2.37100000	7.43400000	2.76700000
H	2.40400000	6.86200000	4.43800000
H	3.34000000	6.01400000	3.19800000
C	1.24200000	7.90600000	-3.85600000
H	0.37900000	8.46100000	-4.23900000
H	1.89100000	8.61300000	-3.33000000
H	1.79700000	7.52700000	-4.72300000
C	6.85400000	-5.29400000	-2.67700000
H	7.85300000	-5.04200000	-3.05500000
H	6.97100000	-6.15900000	-2.01200000
H	6.24500000	-5.60800000	-3.53000000
C	3.42400000	-5.05200000	1.03200000
H	3.84900000	-6.00100000	1.37000000
H	2.33500000	-5.11100000	1.02800000
H	3.77500000	-4.86700000	0.00900000
C	-4.84500000	-1.54900000	3.31100000
H	-5.31000000	-1.08300000	4.19000000
H	-5.12900000	-0.97900000	2.43200000
H	-5.18800000	-2.58500000	3.21900000
C	-8.91900000	-3.33400000	2.83000000
H	-9.87000000	-2.82400000	2.64200000
H	-9.12800000	-4.38400000	3.03300000

H	-8.43300000	-2.90100000	3.71000000
C	-3.66800000	-2.00400000	-3.76600000
H	-4.59900000	-2.34100000	-4.22700000
H	-3.36500000	-1.05700000	-4.21600000
H	-2.89100000	-2.75500000	-3.90500000
O	-0.60500000	-0.00900000	0.09400000
H	-1.00100000	0.24800000	0.94900000
H	0.22100000	-0.50900000	0.28100000
H	-0.07200000	-2.50000000	0.85400000
O	1.34300000	-1.87800000	0.50400000
H	2.02000000	-2.22700000	1.14400000
H	1.80300000	-1.64900000	-0.34900000

Table S11. Cartesian coordinates for ES FeNH

C	-3.09300000	-4.75200000	-1.99100000
O	-1.91300000	-4.65300000	-2.38500000
N	-4.13100000	-4.50800000	-2.81400000
S	4.50500000	-0.80500000	2.11300000
N	4.87500000	2.51200000	1.94700000
C	3.49900000	2.70800000	1.50500000
C	3.19500000	2.48400000	0.01600000
S	3.43900000	0.78500000	-0.56000000
C	2.51000000	2.01400000	2.47100000
O	2.28700000	2.58000000	3.54600000
O	2.66900000	0.75500000	-1.90000000
O	4.95300000	0.63400000	-0.81000000
N	1.84100000	0.87000000	2.10000000
C	0.79300000	0.47300000	3.06800000
C	0.54300000	-1.02100000	2.97100000
O	-0.47300000	-1.56400000	3.49200000
C	-0.48600000	1.32400000	2.98900000
O	-1.02700000	1.44300000	1.66800000
N	1.53000000	-1.67700000	2.37700000
C	1.48800000	-3.13700000	2.35600000
C	2.84100000	-3.65000000	1.85600000
S	3.35300000	-2.73000000	0.34900000
C	1.22100000	-3.70600000	3.76400000
O	2.03200000	-3.62000000	4.66000000
O	2.42700000	-3.57700000	-0.78600000
C	-6.10200000	-4.30300000	2.40500000
C	-5.11300000	-3.40200000	2.72500000
C	-6.51100000	-3.98400000	1.05900000
N	-4.90300000	-2.52100000	1.67200000
C	-5.74000000	-2.86200000	0.63100000
C	-7.44700000	-4.54800000	0.17300000
C	-5.90100000	-2.28900000	-0.63600000
C	-7.60000000	-3.99200000	-1.09600000
C	-6.84000000	-2.86600000	-1.49100000
C	-6.62800000	1.99500000	-1.75000000
C	-5.66700000	2.85500000	-2.30700000
C	-6.34900000	1.45000000	-0.48800000
C	-4.48300000	3.16600000	-1.63700000
C	-5.17100000	1.75900000	0.20400000
C	-4.23100000	2.62300000	-0.37100000
O	-3.05500000	2.97700000	0.22900000
N	3.17800000	-1.55400000	-3.16500000
C	2.46800000	-2.60100000	-3.54800000
N	3.01900000	-3.57400000	-4.30400000
N	1.15900000	-2.67700000	-3.23100000

C	-1.09600000	6.89200000	2.94100000
C	-0.02700000	6.04800000	3.27500000
C	-1.77600000	6.64600000	1.74100000
C	0.37100000	5.00700000	2.43700000
C	-1.40300000	5.59900000	0.89500000
C	-0.31900000	4.79500000	1.24400000
O	0.09200000	3.73500000	0.43600000
C	-1.57800000	5.64100000	-3.17600000
C	-1.36100000	4.42100000	-3.83100000
C	-0.53600000	6.13300000	-2.37300000
C	-0.16200000	3.71200000	-3.69300000
C	0.65800000	5.43000000	-2.20800000
C	0.85100000	4.20400000	-2.86200000
O	2.05100000	3.57200000	-2.66200000
N	8.00100000	-2.78600000	-1.70800000
C	7.37300000	-1.62500000	-1.42500000
N	7.51200000	-0.56800000	-2.25100000
N	6.58600000	-1.50300000	-0.36400000
Fe	2.49800000	-0.60000000	1.07600000
H	-5.07900000	-4.44300000	-2.46300000
H	-3.94000000	-4.18600000	-3.75300000
H	-1.24400000	0.83500000	3.60500000
H	-0.27400000	2.31600000	3.39700000
H	-0.51300000	2.12600000	1.19900000
H	1.16900000	0.62500000	4.09200000
H	-4.56300000	-3.28700000	3.64900000
H	-4.11000000	-1.90000000	1.60900000
H	-8.04700000	-5.40100000	0.47700000
H	-8.33000000	-4.41000000	-1.78300000
H	-5.33700000	-1.41300000	-0.93700000
H	-7.00500000	-2.42700000	-2.47200000
H	-7.06500000	0.77500000	-0.02500000
H	-5.84800000	3.29400000	-3.28600000
H	-4.98600000	1.33300000	1.18700000
H	-3.74300000	3.82500000	-2.08000000
H	-2.82200000	2.42300000	0.99500000
H	2.73000000	-0.74200000	-2.68600000
H	2.55400000	-4.46000000	-4.42500000
H	3.93700000	-3.46700000	-4.70300000
H	0.62200000	-3.48800000	-3.50900000
H	0.88100000	-2.30100000	-2.33200000
H	-2.62300000	7.27000000	1.46600000
H	0.51200000	6.20400000	4.20700000
H	-1.96200000	5.39200000	-0.01000000
H	1.19300000	4.36100000	2.72100000
H	-0.21000000	3.90100000	-0.47100000
H	-0.66300000	7.07600000	-1.84700000
H	-2.13800000	4.01400000	-4.47500000
H	1.44600000	5.81200000	-1.56700000
H	-0.01000000	2.78400000	-4.23700000
H	2.00400000	2.60100000	-2.73100000
H	7.95200000	-3.50800000	-1.00300000
H	6.29100000	-2.30500000	0.17600000
H	5.97500000	-0.66400000	-0.28700000
H	8.29400000	-0.49900000	-2.88400000
H	7.01600000	0.28700000	-2.01800000
H	5.50200000	3.13900000	1.45100000
H	5.18200000	1.55500000	1.79600000
H	3.27100000	3.76800000	1.66800000
H	2.15000000	2.71300000	-0.19600000

H	3.83000000	3.10100000	-0.62700000
H	0.69300000	-3.50000000	1.69700000
H	2.80800000	-4.70400000	1.57000000
H	3.61400000	-3.49500000	2.60700000
H	0.25800000	-4.23200000	3.90800000
C	-7.93400000	1.71200000	-2.46600000
H	-8.34100000	0.73500000	-2.18300000
H	-8.69000000	2.46600000	-2.21600000
H	-7.80600000	1.72500000	-3.55300000
C	-2.88000000	6.40700000	-3.30700000
H	-3.45000000	6.38300000	-2.37000000
H	-2.69900000	7.45900000	-3.55100000
H	-3.51500000	5.98400000	-4.09200000
C	-1.48000000	8.03900000	3.84200000
H	-0.81200000	8.89700000	3.69600000
H	-2.50100000	8.38100000	3.64500000
H	-1.41700000	7.75600000	4.89800000
C	-6.69000000	-5.36500000	3.28900000
H	-7.75300000	-5.17500000	3.48300000
H	-6.61800000	-6.35800000	2.82900000
H	-6.17700000	-5.40600000	4.25500000
C	-3.42300000	-5.12800000	-0.56500000
H	-4.46700000	-5.41400000	-0.42400000
H	-2.76500000	-5.94200000	-0.25400000
H	-3.22400000	-4.26100000	0.07300000
C	4.62600000	-1.43200000	-3.28700000
H	4.93900000	-1.31100000	-4.33300000
H	4.90700000	-0.54200000	-2.73200000
H	5.12500000	-2.30400000	-2.85100000
C	8.77300000	-3.07000000	-2.92300000
H	9.73100000	-2.53800000	-2.93100000
H	8.97700000	-4.14000000	-2.95200000
H	8.19600000	-2.80300000	-3.81400000
C	4.15900000	-0.86200000	3.91200000
H	5.12000000	-0.96200000	4.42100000
H	3.69200000	0.07900000	4.21100000
H	3.51000000	-1.69900000	4.17200000
H	1.45300000	-3.55800000	-0.57100000
O	-0.24100000	-3.56700000	-0.56100000
H	-0.78700000	-4.12900000	-1.15800000
H	-0.78500000	-3.38400000	0.23600000
O	-1.80800000	-3.01200000	1.71500000
H	-2.24300000	-3.73400000	2.18500000
H	-1.33600000	-2.47400000	2.41600000
N	0.84700000	-0.74000000	-0.10700000
C	-0.15600000	-0.59200000	-0.66500000
C	-1.41700000	-0.27700000	-1.36100000
C	-1.31300000	1.20800000	-1.76900000
H	-2.21300000	1.48900000	-2.32300000
H	-1.24600000	1.86100000	-0.90200000
H	-0.44300000	1.36500000	-2.40800000
C	-1.60100000	-1.13900000	-2.62900000
H	-1.61000000	-2.20600000	-2.40100000
H	-2.56900000	-0.87300000	-3.06600000
H	-0.82900000	-0.93100000	-3.37300000
C	-2.58900000	-0.53200000	-0.39000000
H	-3.50200000	-0.15200000	-0.85400000
H	-2.69300000	-1.60100000	-0.19300000
H	-2.42900000	-0.00900000	0.55400000

Table S12. Cartesian coordinates for TS1 FeNH

C	-2.34500000	-5.00300000	-1.89900000
O	-1.10800000	-5.14100000	-1.99900000
N	-3.09500000	-4.56800000	-2.92800000
S	4.24200000	-0.68500000	2.37100000
N	4.81300000	2.25600000	2.09200000
C	3.64900000	2.93900000	1.49200000
C	3.48400000	2.64700000	-0.00100000
S	3.45500000	0.87500000	-0.39300000
C	2.42500000	2.63100000	2.37900000
O	2.13400000	3.48000000	3.23100000
O	2.85100000	0.75400000	-1.77900000
O	4.95300000	0.46100000	-0.40700000
N	1.74200000	1.45700000	2.23700000
C	0.62200000	1.28300000	3.18900000
C	0.38900000	-0.19700000	3.43500000
O	-0.50200000	-0.60100000	4.21900000
C	-0.67800000	1.93200000	2.69100000
O	-1.18900000	1.28400000	1.50500000
N	1.17400000	-0.98500000	2.70700000
C	0.90300000	-2.41900000	2.70900000
C	2.17600000	-3.18300000	2.31600000
S	3.26900000	-2.21900000	1.18500000
C	0.45600000	-2.90700000	4.11200000
O	1.25900000	-3.08200000	5.00800000
O	1.04400000	-2.65900000	-0.61300000
C	-5.73100000	-4.56300000	2.42700000
C	-4.84900000	-3.57200000	2.79700000
C	-6.03600000	-4.32000000	1.03900000
N	-4.60300000	-2.71400000	1.74200000
C	-5.31500000	-3.15000000	0.64800000
C	-6.81600000	-4.99000000	0.07900000
C	-5.37300000	-2.64000000	-0.65300000
C	-6.86200000	-4.49800000	-1.22500000
C	-6.15000000	-3.32800000	-1.58600000
C	-6.68600000	1.44600000	-2.02000000
C	-5.81400000	2.37000000	-2.62200000
C	-6.24000000	0.81100000	-0.85100000
C	-4.56100000	2.66100000	-2.08000000
C	-4.98400000	1.08800000	-0.29200000
C	-4.14500000	2.02300000	-0.90600000
O	-2.90300000	2.35500000	-0.41900000
N	3.62600000	-1.86200000	-2.84900000
C	3.27500000	-3.13800000	-3.14800000
N	4.02400000	-3.76900000	-4.12700000
N	2.33200000	-3.73300000	-2.46800000
C	-1.60200000	6.97900000	2.51900000
C	-0.28400000	6.49800000	2.47600000
C	-2.53700000	6.38400000	1.66200000
C	0.08600000	5.45700000	1.62600000
C	-2.18600000	5.33500000	0.80500000
C	-0.87100000	4.86000000	0.79900000
O	-0.46900000	3.80000000	0.02400000
C	-2.01800000	5.52000000	-3.37100000
C	-1.70600000	4.26900000	-3.91600000
C	-1.03200000	6.14300000	-2.58400000
C	-0.47300000	3.64600000	-3.67500000
C	0.19700000	5.54000000	-2.33600000
C	0.47600000	4.27400000	-2.86600000
O	1.69100000	3.71100000	-2.54400000

N	7.89900000	-2.64000000	-2.14100000
C	7.17600000	-1.66800000	-1.55000000
N	7.44600000	-0.37300000	-1.76700000
N	6.15500000	-1.97400000	-0.74400000
Fe	2.22700000	-0.12300000	1.32400000
H	-4.07900000	-4.36700000	-2.80100000
H	-2.64400000	-4.31500000	-3.79600000
H	-1.44800000	1.81800000	3.45600000
H	-0.50600000	2.99100000	2.49500000
H	-0.46400000	1.22400000	0.85500000
H	0.86600000	1.76800000	4.14100000
H	-4.38400000	-3.39800000	3.75900000
H	-3.86800000	-2.00400000	1.75800000
H	-7.37700000	-5.88100000	0.35000000
H	-7.46900000	-5.00300000	-1.97000000
H	-4.84700000	-1.72900000	-0.91600000
H	-6.24100000	-2.93600000	-2.59600000
H	-6.87800000	0.08100000	-0.36100000
H	-6.12200000	2.87900000	-3.53200000
H	-4.67000000	0.58700000	0.62000000
H	-3.89600000	3.38000000	-2.54600000
H	-2.66800000	1.91500000	0.42000000
H	2.98700000	-1.29400000	-2.29800000
H	3.69100000	-4.67900000	-4.41400000
H	4.36000000	-3.20600000	-4.89700000
H	2.05000000	-4.64700000	-2.80600000
H	1.62000000	-3.13900000	-1.44300000
H	-3.56600000	6.73600000	1.66500000
H	0.46700000	6.93200000	3.13100000
H	-2.92800000	4.87700000	0.15900000
H	1.09700000	5.06800000	1.64200000
H	-1.22800000	3.47700000	-0.49100000
H	-1.24000000	7.11100000	-2.13500000
H	-2.43400000	3.76000000	-4.54300000
H	0.93000000	6.01400000	-1.69200000
H	-0.26000000	2.67600000	-4.11400000
H	1.71900000	2.76600000	-2.75400000
H	7.74200000	-3.58100000	-1.80900000
H	5.71700000	-2.88100000	-0.80900000
H	5.55600000	-1.18200000	-0.43900000
H	8.31200000	-0.07100000	-2.18600000
H	6.79100000	0.32300000	-1.40900000
H	5.11000000	2.75600000	2.92500000
H	5.59400000	2.20700000	1.44400000
H	3.74100000	4.03100000	1.55800000
H	2.54500000	3.06400000	-0.37900000
H	4.31300000	3.05900000	-0.58700000
H	0.09800000	-2.62100000	1.99300000
H	1.94700000	-4.09100000	1.75200000
H	2.77200000	-3.44300000	3.19300000
H	-0.61200000	-3.15600000	4.23000000
C	-8.06500000	1.18000000	-2.59800000
H	-8.46800000	0.22600000	-2.24300000
H	-8.77200000	1.96700000	-2.30700000
H	-8.03900000	1.15300000	-3.69300000
C	-3.35800000	6.19100000	-3.61000000
H	-3.83900000	6.45900000	-2.66200000
H	-3.24100000	7.11300000	-4.19000000
H	-4.04000000	5.53500000	-4.16000000
C	-1.98900000	8.07800000	3.48300000

H	-1.36300000	8.96800000	3.34900000
H	-3.03200000	8.38100000	3.34700000
H	-1.87000000	7.75200000	4.52300000
C	-6.26800000	-5.66100000	3.29400000
H	-7.35800000	-5.59300000	3.39900000
H	-6.04800000	-6.65200000	2.87600000
H	-5.83500000	-5.61900000	4.29800000
C	-3.07000000	-5.28800000	-0.60800000
H	-4.14900000	-5.38400000	-0.73300000
H	-2.67600000	-6.21100000	-0.17500000
H	-2.89700000	-4.46700000	0.09500000
C	4.67200000	-1.11100000	-3.51800000
H	4.38400000	-0.80400000	-4.53400000
H	4.84700000	-0.20800000	-2.93700000
H	5.59200000	-1.69700000	-3.58000000
C	8.92600000	-2.45100000	-3.17300000
H	9.85900000	-2.05800000	-2.75300000
H	9.13300000	-3.42000000	-3.62700000
H	8.55900000	-1.77700000	-3.95300000
C	3.97100000	-1.08000000	4.13300000
H	4.21700000	-0.14300000	4.64400000
H	2.93600000	-1.34700000	4.33400000
H	4.66100000	-1.86600000	4.44400000
H	0.46900000	-3.34700000	-0.19600000
O	-0.00500000	-4.88800000	0.42000000
H	-0.39000000	-5.17300000	-0.45300000
H	-0.70400000	-5.02100000	1.07000000
O	-2.24800000	-1.21700000	2.10400000
H	-1.92000000	-1.20700000	3.02200000
H	-1.92100000	-0.35500000	1.77200000
N	0.80700000	-0.26300000	0.09500000
C	0.33700000	-1.02200000	-0.70700000
C	-0.69100000	-0.98900000	-1.81400000
C	-0.76000000	0.47900000	-2.28600000
H	-1.50200000	0.56900000	-3.08500000
H	-1.05900000	1.15500000	-1.48600000
H	0.21600000	0.79600000	-2.66300000
C	-0.30300000	-1.87800000	-3.00900000
H	-0.30800000	-2.93700000	-2.75500000
H	-1.02900000	-1.70700000	-3.81100000
H	0.68700000	-1.61500000	-3.39400000
C	-2.05200000	-1.43300000	-1.24600000
H	-2.78800000	-1.44500000	-2.05500000
H	-2.00400000	-2.42800000	-0.80000000
H	-2.40400000	-0.74100000	-0.47800000

Table S13. Cartesian coordinates for INT1 FeNH

C	-2.74600000	-4.58600000	-1.91300000
O	-1.52000000	-4.47700000	-1.74000000
N	-3.43500000	-3.77300000	-2.74700000
S	4.97400000	0.03400000	2.21600000
N	4.86800000	2.63700000	1.87800000
C	3.53200000	3.05600000	1.43000000
C	3.31800000	2.85500000	-0.07000000
S	3.83800000	1.22100000	-0.68400000
C	2.49900000	2.39300000	2.36700000
O	2.16900000	3.05800000	3.36900000
O	3.26600000	1.10900000	-2.07800000
O	5.38100000	1.30700000	-0.71700000

N	2.04800000	1.15400000	2.09900000
C	1.14000000	0.60500000	3.11800000
C	1.16500000	-0.90700000	3.05600000
O	0.41600000	-1.61300000	3.76600000
C	-0.29600000	1.15000000	3.09000000
O	-1.01100000	0.82200000	1.89200000
N	2.09700000	-1.42400000	2.23900000
C	2.24200000	-2.88300000	2.22100000
C	3.56600000	-3.23100000	1.53500000
S	3.86600000	-2.14900000	0.07500000
C	2.26500000	-3.43500000	3.66200000
O	3.12800000	-3.12400000	4.45900000
O	0.65600000	-2.31700000	-0.34300000
C	-5.93900000	-5.08800000	2.32400000
C	-4.94700000	-4.48700000	3.06500000
C	-6.08600000	-4.28400000	1.13500000
N	-4.46900000	-3.36000000	2.41900000
C	-5.15300000	-3.20600000	1.23500000
C	-6.91400000	-4.36200000	0.00000000
C	-5.05300000	-2.21400000	0.25100000
C	-6.81200000	-3.37900000	-0.98500000
C	-5.89700000	-2.30800000	-0.85400000
C	-7.22000000	1.19600000	-2.02100000
C	-6.10800000	1.63000000	-2.76200000
C	-7.04900000	1.02200000	-0.64000000
C	-4.87500000	1.89100000	-2.15500000
C	-5.82600000	1.27600000	-0.01600000
C	-4.73900000	1.71400000	-0.77600000
O	-3.50600000	1.98000000	-0.21500000
N	3.04300000	-1.82000000	-3.07700000
C	2.66200000	-3.08600000	-3.25200000
N	3.15300000	-3.77700000	-4.31400000
N	1.85200000	-3.69800000	-2.39100000
C	-2.14800000	6.29900000	2.95200000
C	-0.87800000	5.70500000	3.00800000
C	-3.03900000	5.84500000	1.97200000
C	-0.50800000	4.69000000	2.12600000
C	-2.69000000	4.82400000	1.08200000
C	-1.42300000	4.24400000	1.17200000
O	-1.02800000	3.21000000	0.34300000
C	-2.52100000	5.43800000	-3.03900000
C	-2.02300000	4.37900000	-3.81000000
C	-1.68900000	5.95000000	-2.03000000
C	-0.74400000	3.85300000	-3.59500000
C	-0.41400000	5.43500000	-1.79900000
C	0.06300000	4.38000000	-2.58300000
O	1.32700000	3.91400000	-2.30400000
N	7.95100000	-2.25500000	-2.16000000
C	7.35400000	-1.11300000	-1.75800000
N	7.20500000	-0.07400000	-2.59600000
N	6.88500000	-1.00300000	-0.51500000
Fe	2.85000000	-0.27600000	0.95800000
H	-4.44100000	-3.71600000	-2.65900000
H	-2.95700000	-2.96100000	-3.11400000
H	-0.83800000	0.68800000	3.92100000
H	-0.27500000	2.22900000	3.24400000
H	-0.85700000	1.52900000	1.23800000
H	1.51400000	0.86300000	4.12100000
H	-4.54300000	-4.77700000	4.02600000
H	-3.72300000	-2.74900000	2.74400000

H	-7.63400000	-5.17000000	-0.10000000
H	-7.46600000	-3.41600000	-1.85200000
H	-4.36300000	-1.38500000	0.35600000
H	-5.87100000	-1.52500000	-1.60600000
H	-7.88100000	0.66900000	-0.03600000
H	-6.20600000	1.77800000	-3.83500000
H	-5.71900000	1.12400000	1.05600000
H	-4.02500000	2.24400000	-2.73100000
H	-3.52900000	1.85100000	0.74400000
H	2.66300000	-1.29700000	-2.29100000
H	2.86500000	-4.73400000	-4.45300000
H	3.48700000	-3.28400000	-5.12800000
H	1.28000000	-4.47600000	-2.69100000
H	1.52000000	-3.20600000	-1.54700000
H	-4.02900000	6.28900000	1.90000000
H	-0.16300000	6.03100000	3.75800000
H	-3.39200000	4.48500000	0.32700000
H	0.47400000	4.23900000	2.19600000
H	-1.71700000	3.05000000	-0.32300000
H	-2.04800000	6.75900000	-1.39800000
H	-2.63800000	3.96000000	-4.60400000
H	0.20600000	5.82000000	-0.99600000
H	-0.37900000	3.03400000	-4.21100000
H	1.50700000	3.07800000	-2.75700000
H	8.18500000	-2.90900000	-1.42600000
H	6.52500000	-1.82500000	-0.04000000
H	6.33900000	-0.14700000	-0.32200000
H	7.78500000	0.03900000	-3.41200000
H	6.63600000	0.71000000	-2.26900000
H	5.13600000	3.06100000	2.76100000
H	5.59700000	2.72200000	1.17600000
H	3.39300000	4.13000000	1.61000000
H	2.26100000	2.94500000	-0.32400000
H	3.88800000	3.57700000	-0.66500000
H	1.39700000	-3.33000000	1.69400000
H	3.57100000	-4.27300000	1.20200000
H	4.38500000	-3.08300000	2.24300000
H	1.48500000	-4.17600000	3.91400000
C	-8.56700000	0.94800000	-2.67100000
H	-9.03100000	0.03400000	-2.28500000
H	-9.25800000	1.77500000	-2.47100000
H	-8.47600000	0.85000000	-3.75700000
C	-3.88300000	6.04300000	-3.30500000
H	-4.36600000	6.36100000	-2.37500000
H	-3.80400000	6.92600000	-3.95100000
H	-4.54900000	5.33000000	-3.80200000
C	-2.52400000	7.40300000	3.90900000
H	-2.05700000	8.35400000	3.62400000
H	-3.60600000	7.56300000	3.93300000
H	-2.19600000	7.17400000	4.92900000
C	-6.73700000	-6.30500000	2.69500000
H	-7.80400000	-6.06800000	2.79200000
H	-6.64900000	-7.09400000	1.93700000
H	-6.40000000	-6.72100000	3.64900000
C	-3.54100000	-5.62600000	-1.16700000
H	-4.45300000	-5.92400000	-1.69000000
H	-2.90900000	-6.49700000	-0.98700000
H	-3.84900000	-5.21000000	-0.20000000
C	4.18100000	-1.20700000	-3.75500000
H	3.98200000	-1.07000000	-4.82500000

H	4.32100000	-0.22500000	-3.31000000
H	5.08000000	-1.81800000	-3.62400000
C	8.44100000	-2.54800000	-3.50900000
H	9.37000000	-2.01100000	-3.73500000
H	8.63400000	-3.61900000	-3.57600000
H	7.68000000	-2.29000000	-4.25100000
C	4.51100000	0.07500000	3.97400000
H	5.41600000	0.16500000	4.57800000
H	3.84600000	0.91900000	4.16800000
H	3.99700000	-0.86200000	4.21200000
H	-0.14000000	-2.78500000	0.09600000
O	-1.01000000	-3.70200000	0.92100000
H	-1.47800000	-4.20100000	0.23200000
H	-1.62100000	-3.12900000	1.43400000
O	-2.05500000	-1.75000000	2.53600000
H	-1.30200000	-1.82000000	3.16300000
H	-1.91900000	-0.86000000	2.15300000
N	1.41900000	-0.14700000	-0.25900000
C	0.56200000	-0.99600000	-0.65000000
C	-0.56900000	-0.57300000	-1.61100000
C	-0.42100000	0.90700000	-1.98300000
H	-1.29000000	1.23900000	-2.55700000
H	-0.31600000	1.53900000	-1.10300000
H	0.48000000	1.05500000	-2.58600000
C	-0.49400000	-1.41800000	-2.90300000
H	-0.60300000	-2.48400000	-2.69300000
H	-1.30100000	-1.10000000	-3.57400000
H	0.44600000	-1.24800000	-3.43700000
C	-1.93100000	-0.82700000	-0.93300000
H	-2.73800000	-0.53200000	-1.61100000
H	-2.06800000	-1.88100000	-0.68500000
H	-2.03300000	-0.23100000	-0.02700000

Table S14. Cartesian coordinates for TS2 FeNH

C	2.77500000	-4.95100000	2.15000000
O	1.62000000	-4.54900000	1.93500000
N	3.39200000	-4.68900000	3.32900000
S	-5.55800000	0.49900000	-2.21200000
N	-4.95100000	3.05300000	-1.90000000
C	-3.56100000	3.33500000	-1.51600000
C	-3.32500000	3.16100000	-0.00700000
S	-4.04100000	1.64100000	0.74300000
C	-2.58500000	2.55200000	-2.43400000
O	-1.98200000	3.20900000	-3.30100000
O	-3.43900000	1.59500000	2.13700000
O	-5.56100000	1.97700000	0.78800000
N	-2.41800000	1.22000000	-2.23900000
C	-1.45600000	0.58200000	-3.11400000
C	-1.64500000	-0.94800000	-3.17700000
O	-0.82200000	-1.66500000	-3.78900000
C	0.01200000	0.95300000	-2.65800000
O	0.26200000	0.73000000	-1.32100000
N	-2.73500000	-1.41400000	-2.55000000
C	-2.94200000	-2.85900000	-2.49700000
C	-4.26900000	-3.15200000	-1.76200000
S	-4.49500000	-2.15100000	-0.21500000
C	-3.01900000	-3.45500000	-3.91400000
O	-3.83900000	-3.08800000	-4.73100000
O	-1.28600000	-2.41700000	0.90100000

C	5.67900000	-5.43800000	-2.46100000
C	4.49300000	-5.00700000	-3.00900000
C	5.97200000	-4.52400000	-1.38600000
N	4.03100000	-3.87900000	-2.34900000
C	4.91600000	-3.56500000	-1.33900000
C	7.01400000	-4.45000000	-0.44200000
C	4.88600000	-2.55200000	-0.36900000
C	6.98000000	-3.44800000	0.52400000
C	5.92500000	-2.51000000	0.56000000
C	7.83700000	0.72500000	1.72400000
C	6.65500000	1.01500000	2.41900000
C	7.79000000	0.73500000	0.32200000
C	5.45900000	1.28700000	1.74900000
C	6.60600000	1.00400000	-0.36500000
C	5.43900000	1.26700000	0.35500000
O	4.22800000	1.49800000	-0.27200000
N	-3.41400000	-1.89400000	3.20800000
C	-3.15800000	-3.15200000	3.55800000
N	-3.29800000	-3.54500000	4.84400000
N	-2.80200000	-4.02500000	2.61000000
C	2.90800000	6.20900000	-2.92000000
C	1.61400000	5.66400000	-2.96600000
C	3.83800000	5.63800000	-2.03700000
C	1.25200000	4.59100000	-2.16200000
C	3.49900000	4.56100000	-1.22300000
C	2.19900000	4.03300000	-1.28400000
O	1.79800000	2.98100000	-0.53000000
C	3.43100000	5.45700000	3.06300000
C	2.77800000	4.45500000	3.79300000
C	2.68900000	6.11100000	2.06600000
C	1.44600000	4.10700000	3.53600000
C	1.36300000	5.77800000	1.79600000
C	0.73600000	4.76500000	2.52700000
O	-0.56500000	4.45900000	2.20300000
N	-7.83800000	-1.55100000	2.64900000
C	-7.34400000	-0.41200000	2.12500000
N	-7.11800000	0.65900000	2.90600000
N	-7.06200000	-0.33200000	0.82500000
Fe	-3.35300000	-0.23200000	-1.08100000
H	4.36300000	-4.91900000	3.47500000
H	2.93800000	-4.07300000	3.98900000
H	0.70600000	0.39200000	-3.29800000
H	0.08500000	2.03000000	-2.88600000
H	-0.67900000	0.47300000	-0.61000000
H	-1.52500000	0.97000000	-4.14000000
H	3.93100000	-5.42000000	-3.83700000
H	3.12400000	-3.44800000	-2.50500000
H	7.83000000	-5.16600000	-0.46500000
H	7.78100000	-3.37700000	1.25500000
H	4.08300000	-1.82200000	-0.31900000
H	5.93100000	-1.72800000	1.31200000
H	8.68900000	0.51200000	-0.24700000
H	6.66200000	1.02400000	3.50600000
H	6.58900000	0.98400000	-1.45200000
H	4.54800000	1.50800000	2.29600000
H	4.33200000	1.42600000	-1.23100000
H	-3.38300000	-1.69900000	2.21000000
H	-2.91700000	-4.42700000	5.14900000
H	-3.51100000	-2.86800000	5.56100000
H	-2.49300000	-4.95300000	2.85300000

H	-2.42900000	-3.59600000	1.75500000
H	4.84400000	6.04600000	-1.98400000
H	0.88000000	6.08200000	-3.65100000
H	4.22200000	4.13900000	-0.53200000
H	0.25900000	4.15900000	-2.21900000
H	2.56100000	2.58300000	-0.07300000
H	3.16100000	6.89300000	1.47500000
H	3.31400000	3.93200000	4.58300000
H	0.81100000	6.28100000	1.00800000
H	0.96700000	3.32400000	4.12000000
H	-0.88400000	3.72100000	2.73900000
H	-8.11100000	-2.26400000	1.98800000
H	-6.74600000	-1.16400000	0.32500000
H	-6.60700000	0.55200000	0.53800000
H	-7.58700000	0.77900000	3.78900000
H	-6.61800000	1.44700000	2.48600000
H	-5.24100000	3.54100000	-2.74100000
H	-5.62200000	3.16200000	-1.14400000
H	-3.32100000	4.38300000	-1.73800000
H	-2.25500000	3.14400000	0.21700000
H	-3.78700000	3.97600000	0.56200000
H	-2.10800000	-3.33700000	-1.96800000
H	-4.32600000	-4.20900000	-1.49000000
H	-5.10500000	-2.92200000	-2.42500000
H	-2.31600000	-4.27900000	-4.14300000
C	9.12400000	0.42000000	2.45300000
H	9.66400000	-0.40700000	1.98100000
H	9.79700000	1.28700000	2.45600000
H	8.93700000	0.14800000	3.49600000
C	4.87000000	5.84200000	3.34600000
H	5.42100000	6.03500000	2.42000000
H	4.92700000	6.75400000	3.95300000
H	5.39500000	5.05100000	3.89100000
C	3.28100000	7.37900000	-3.78600000
H	2.81900000	8.30700000	-3.42200000
H	4.36200000	7.53700000	-3.80500000
H	2.94100000	7.23800000	-4.81800000
C	6.50700000	-6.61700000	-2.87500000
H	7.51500000	-6.31600000	-3.18700000
H	6.62600000	-7.33500000	-2.05300000
H	6.04700000	-7.14900000	-3.71300000
C	3.55500000	-5.74500000	1.14000000
H	4.26600000	-6.42600000	1.61700000
H	2.86300000	-6.32500000	0.52700000
H	4.12700000	-5.08700000	0.48000000
C	-3.68400000	-0.79200000	4.12400000
H	-2.91100000	-0.73200000	4.89800000
H	-3.65900000	0.13400000	3.54200000
H	-4.66700000	-0.89600000	4.59900000
C	-8.04100000	-1.80900000	4.07000000
H	-8.92200000	-1.29100000	4.46500000
H	-8.19000000	-2.88100000	4.20700000
H	-7.15900000	-1.50900000	4.64400000
C	-5.22100000	0.38000000	-3.99900000
H	-6.10700000	0.69100000	-4.55500000
H	-4.37000000	1.01400000	-4.25900000
H	-4.96700000	-0.65500000	-4.24300000
H	-0.38900000	-3.15100000	0.30200000
O	0.68900000	-3.66400000	-0.72000000
H	1.34400000	-3.96300000	-0.04700000

H	1.17600000	-3.22400000	-1.45600000
O	1.34800000	-1.87100000	-2.06200000
H	0.70300000	-2.05200000	-2.78700000
H	1.28600000	-0.91300000	-1.93800000
N	-1.51500000	-0.31400000	-0.01900000
C	-0.86800000	-1.12100000	0.72600000
C	0.38000000	-0.70500000	1.55600000
C	0.39400000	0.82300000	1.75900000
H	1.21000000	1.08800000	2.44000000
H	0.56000000	1.35600000	0.82100000
H	-0.55000000	1.16700000	2.19500000
C	0.30600000	-1.38700000	2.94300000
H	0.35500000	-2.47400000	2.86800000
H	1.15400000	-1.04400000	3.54700000
H	-0.60900000	-1.10300000	3.47500000
C	1.69000000	-1.13000000	0.85600000
H	2.54400000	-0.83600000	1.47800000
H	1.74400000	-2.21000000	0.71700000
H	1.78700000	-0.60900000	-0.09900000

Table S15. Cartesian coordinates for INT2 FeNH

C	-1.00400000	4.98600000	2.30400000
O	0.15100000	4.70000000	2.67500000
N	-2.08800000	4.57600000	2.98600000
S	3.75100000	0.14800000	-2.80800000
N	3.41800000	-2.56700000	-3.25400000
C	2.20400000	-3.06400000	-2.57300000
C	2.35600000	-3.16600000	-1.05600000
S	2.90100000	-1.63600000	-0.22500000
C	1.00300000	-2.22900000	-3.05700000
O	0.33900000	-2.70400000	-3.98800000
O	2.54200000	-1.84500000	1.22300000
O	4.44900000	-1.65500000	-0.44400000
N	0.73100000	-1.02200000	-2.48900000
C	-0.40300000	-0.31000000	-3.11800000
C	-0.22200000	1.18600000	-2.93200000
O	-1.11300000	1.99800000	-3.26100000
C	-1.77500000	-0.78000000	-2.61900000
O	-1.82400000	-0.83600000	-1.18600000
N	0.98000000	1.51000000	-2.46200000
C	1.22800000	2.92700000	-2.22700000
C	2.73800000	3.16600000	-2.08700000
S	3.61100000	1.73500000	-1.30900000
C	0.72900000	3.76500000	-3.42400000
O	1.25200000	3.68700000	-4.51600000
O	2.07600000	1.18800000	1.56700000
C	-3.93400000	6.15800000	-2.17400000
C	-3.09800000	5.46300000	-3.01500000
C	-4.15400000	5.30100000	-1.03700000
N	-2.77400000	4.23000000	-2.47300000
C	-3.41800000	4.09900000	-1.26800000
C	-4.90100000	5.42400000	0.14600000
C	-3.42700000	3.03300000	-0.36300000
C	-4.90800000	4.36600000	1.05400000
C	-4.17900000	3.18200000	0.79900000
C	-6.63000000	0.29400000	2.62100000
C	-5.87100000	-0.65200000	3.32600000
C	-6.27500000	0.53600000	1.28700000
C	-4.80000000	-1.32600000	2.73500000

C	-5.21000000	-0.12900000	0.67800000
C	-4.46100000	-1.06200000	1.40400000
O	-3.39600000	-1.72800000	0.85900000
N	4.09300000	0.20600000	3.22700000
C	4.13100000	0.66500000	4.47400000
N	4.91500000	0.03300000	5.39300000
N	3.45200000	1.75600000	4.81500000
C	-3.97600000	-5.85200000	-3.11000000
C	-2.66600000	-5.54300000	-3.50300000
C	-4.44700000	-5.28800000	-1.91700000
C	-1.83400000	-4.74200000	-2.72100000
C	-3.63800000	-4.47100000	-1.12500000
C	-2.32300000	-4.22400000	-1.52200000
O	-1.46000000	-3.46200000	-0.73600000
C	-3.37500000	-5.08200000	3.00000000
C	-2.52900000	-4.23600000	3.73200000
C	-2.82500000	-5.71600000	1.87800000
C	-1.21400000	-3.98400000	3.33500000
C	-1.51300000	-5.47500000	1.46600000
C	-0.70700000	-4.57300000	2.17000000
O	0.54600000	-4.32200000	1.67000000
N	8.51800000	-0.13400000	1.01200000
C	7.40600000	-0.67500000	0.47600000
N	7.10100000	-1.96600000	0.65800000
N	6.56200000	0.07200000	-0.24400000
Fe	1.86300000	0.09700000	-1.44300000
H	-3.00800000	4.63900000	2.56600000
H	-1.94800000	3.92500000	3.74500000
H	-2.52300000	-0.05800000	-2.95500000
H	-1.99700000	-1.76600000	-3.03600000
H	-1.47000000	-1.72600000	-0.96300000
H	-0.39200000	-0.49300000	-4.20100000
H	-2.70600000	5.75400000	-3.98100000
H	-2.21200000	3.49800000	-2.90200000
H	-5.47900000	6.32400000	0.34100000
H	-5.51300000	4.43300000	1.95500000
H	-2.87800000	2.12100000	-0.57300000
H	-4.22500000	2.36400000	1.50800000
H	-6.82800000	1.27300000	0.71000000
H	-6.11700000	-0.86500000	4.36400000
H	-4.94700000	0.09300000	-0.35300000
H	-4.20900000	-2.04800000	3.28800000
H	-3.15600000	-1.32700000	-0.00100000
H	3.32100000	0.53000000	2.60600000
H	4.94700000	0.39200000	6.33400000
H	5.07000000	-0.95900000	5.29700000
H	3.38400000	2.01800000	5.78600000
H	3.16400000	2.46500000	4.09400000
H	-5.46700000	-5.48700000	-1.59700000
H	-2.28100000	-5.94300000	-4.43800000
H	-4.01300000	-4.02400000	-0.21000000
H	-0.82900000	-4.50000000	-3.04300000
H	-1.75800000	-3.53100000	0.18700000
H	-3.43800000	-6.39200000	1.28900000
H	-2.90400000	-3.75600000	4.63300000
H	-1.11400000	-5.95100000	0.57700000
H	-0.58300000	-3.32600000	3.92400000
H	0.91200000	-3.49000000	2.00700000
H	8.76700000	0.79000000	0.68900000
H	6.57600000	1.07800000	-0.17100000

H	5.67000000	-0.37800000	-0.52300000
H	7.78900000	-2.63200000	0.96800000
H	6.18400000	-2.28600000	0.33000000
H	3.43300000	-2.90700000	-4.21100000
H	4.25900000	-2.88800000	-2.78200000
H	1.95500000	-4.08000000	-2.90800000
H	1.39900000	-3.43200000	-0.59700000
H	3.10200000	-3.91700000	-0.77400000
H	0.73100000	3.25900000	-1.31200000
H	2.93900000	3.99100000	-1.40200000
H	3.21200000	3.35200000	-3.05200000
H	-0.08000000	4.48800000	-3.21100000
C	-7.79900000	1.02000000	3.25900000
H	-7.83400000	2.06900000	2.94600000
H	-8.75500000	0.56500000	2.97300000
H	-7.73800000	0.99400000	4.35100000
C	-4.83200000	-5.26200000	3.37600000
H	-5.25000000	-6.17000000	2.93100000
H	-4.96200000	-5.32400000	4.46100000
H	-5.42800000	-4.41100000	3.02400000
C	-4.83300000	-6.78800000	-3.93300000
H	-4.62400000	-7.83500000	-3.68400000
H	-5.89900000	-6.61700000	-3.75600000
H	-4.64500000	-6.66500000	-5.00400000
C	-4.53100000	7.51800000	-2.38400000
H	-5.62700000	7.47500000	-2.40100000
H	-4.25000000	8.21200000	-1.58300000
H	-4.20200000	7.95400000	-3.33200000
C	-1.24500000	5.81200000	1.06200000
H	-0.88800000	5.24400000	0.19700000
H	-2.29400000	6.05700000	0.89800000
H	-0.64900000	6.72500000	1.12900000
C	5.03700000	-0.79100000	2.73400000
H	4.83900000	-1.78900000	3.14500000
H	4.89100000	-0.86900000	1.66400000
H	6.06300000	-0.49300000	2.97100000
C	9.44000000	-0.79000000	1.94800000
H	10.10000000	-1.50300000	1.44200000
H	10.05300000	-0.02100000	2.41700000
H	8.87500000	-1.30500000	2.73000000
C	3.32000000	0.87100000	-4.42700000
H	4.18400000	1.38600000	-4.84700000
H	3.06900000	-0.00300000	-5.03400000
H	2.44900000	1.52300000	-4.35400000
O	2.93400000	3.80000000	3.19000000
H	2.10200000	4.24000000	3.43100000
H	2.85200000	3.75400000	2.21600000
O	1.76200000	3.86900000	0.65400000
H	1.12700000	4.27400000	1.28200000
H	1.74900000	2.91400000	0.86600000
N	0.63800000	0.41400000	-0.00600000
C	0.92800000	0.74000000	1.24000000
C	-0.17900000	0.58500000	2.31200000
C	-0.34200000	-0.91800000	2.60400000
H	-1.10800000	-1.06600000	3.37000000
H	-0.66400000	-1.46800000	1.71600000
H	0.60300000	-1.34500000	2.95300000
C	0.23200000	1.31600000	3.59800000
H	0.40500000	2.38000000	3.41600000
H	-0.56300000	1.21300000	4.34600000

H	1.14700000	0.89400000	4.01800000
C	-1.52100000	1.15200000	1.81100000
H	-2.27300000	1.07700000	2.60400000
H	-1.43300000	2.20200000	1.52500000
H	-1.91100000	0.59200000	0.96000000
H	-0.32100000	0.14200000	-0.21200000

Table S16. Cartesian coordinates for TS3 FeNH

C	-0.90500000	4.49000000	2.80000000
O	0.07400000	3.89400000	3.27200000
N	-2.17000000	4.15600000	3.14600000
S	3.87000000	0.15200000	-2.92700000
N	3.29700000	-2.64000000	-3.35600000
C	2.07400000	-2.95600000	-2.59200000
C	2.29700000	-3.00400000	-1.08800000
S	2.96700000	-1.49300000	-0.31700000
C	0.97000000	-2.00100000	-3.08000000
O	0.32100000	-2.39700000	-4.06000000
O	2.61700000	-1.68200000	1.13500000
O	4.50600000	-1.58500000	-0.53200000
N	0.75800000	-0.78800000	-2.48700000
C	-0.30100000	-0.01100000	-3.17800000
C	-0.05900000	1.47100000	-2.98100000
O	-0.89400000	2.33700000	-3.32400000
C	-1.73600000	-0.44300000	-2.84100000
O	-1.96300000	-0.53600000	-1.43300000
N	1.16600000	1.74300000	-2.52400000
C	1.56900000	3.14600000	-2.47200000
C	3.10200000	3.20400000	-2.46400000
S	3.82300000	1.81500000	-1.48400000
C	1.05600000	3.91400000	-3.71400000
O	1.52200000	3.71500000	-4.81700000
O	2.04100000	0.79900000	2.18000000
C	-3.42800000	6.44400000	-1.52000000
C	-2.63200000	5.77200000	-2.42200000
C	-3.87600000	5.45400000	-0.57400000
N	-2.55900000	4.42700000	-2.10000000
C	-3.32400000	4.20100000	-0.97900000
C	-4.70600000	5.49900000	0.56200000
C	-3.60300000	3.00900000	-0.29700000
C	-4.97200000	4.31500000	1.25100000
C	-4.42700000	3.08200000	0.82200000
C	-6.68800000	0.47800000	2.69600000
C	-5.62700000	-0.15700000	3.35900000
C	-6.71400000	0.40300000	1.29400000
C	-4.61500000	-0.81900000	2.65600000
C	-5.71300000	-0.25500000	0.57700000
C	-4.64900000	-0.85500000	1.25900000
O	-3.62400000	-1.49300000	0.60300000
N	4.24700000	-0.29400000	3.31500000
C	4.32200000	0.36000000	4.47700000
N	5.08200000	-0.17500000	5.48400000
N	3.72700000	1.53300000	4.65400000
C	-4.32100000	-5.34400000	-3.40400000
C	-3.01400000	-5.00100000	-3.78500000
C	-4.75700000	-4.92800000	-2.13800000
C	-2.15800000	-4.30100000	-2.93200000
C	-3.92100000	-4.21700000	-1.27100000
C	-2.61400000	-3.92500000	-1.66800000

O	-1.73200000	-3.24600000	-0.83500000
C	-3.84700000	-5.06300000	2.73700000
C	-3.02100000	-4.19100000	3.45900000
C	-3.26500000	-5.76800000	1.67100000
C	-1.67200000	-4.01600000	3.13100000
C	-1.92400000	-5.60100000	1.32300000
C	-1.12200000	-4.71500000	2.05000000
O	0.18700000	-4.57300000	1.64800000
N	8.43200000	-0.85500000	1.76800000
C	7.35300000	-1.10000000	0.99500000
N	7.03700000	-2.34500000	0.61700000
N	6.55400000	-0.10200000	0.59900000
Fe	2.03200000	0.31300000	-1.52800000
H	-2.96800000	4.48500000	2.61400000
H	-2.30300000	3.33600000	3.72100000
H	-2.41300000	0.31800000	-3.23700000
H	-1.94500000	-1.40200000	-3.32200000
H	-1.63900000	-1.41800000	-1.16400000
H	-0.20100000	-0.16200000	-4.26300000
H	-2.12200000	6.15300000	-3.29700000
H	-2.04300000	3.71300000	-2.61100000
H	-5.14600000	6.43800000	0.88500000
H	-5.63000000	4.33100000	2.11700000
H	-3.20500000	2.06100000	-0.64300000
H	-4.66600000	2.17700000	1.36800000
H	-7.52300000	0.88300000	0.74800000
H	-5.58200000	-0.12800000	4.44500000
H	-5.74600000	-0.29100000	-0.50900000
H	-3.79400000	-1.30400000	3.17400000
H	-3.44500000	-1.07100000	-0.25900000
H	3.48300000	-0.01600000	2.68400000
H	5.08600000	0.29400000	6.37800000
H	5.17900000	-1.17800000	5.53200000
H	3.76900000	1.96400000	5.56700000
H	3.31200000	2.15300000	3.89300000
H	-5.77000000	-5.15700000	-1.81700000
H	-2.65100000	-5.28700000	-4.76900000
H	-4.27500000	-3.89200000	-0.29900000
H	-1.15700000	-4.03000000	-3.24900000
H	-2.10700000	-3.21900000	0.06000000
H	-3.87600000	-6.44500000	1.08000000
H	-3.43600000	-3.63000000	4.29200000
H	-1.49700000	-6.12500000	0.47500000
H	-1.05700000	-3.33100000	3.70700000
H	0.59800000	-3.81800000	2.08900000
H	8.69500000	0.11400000	1.87500000
H	6.57100000	0.79100000	1.07000000
H	5.69200000	-0.36900000	0.09200000
H	7.68600000	-3.11000000	0.70700000
H	6.14500000	-2.48100000	0.13000000
H	3.22200000	-3.01600000	-4.29700000
H	4.12600000	-3.02900000	-2.91500000
H	1.68600000	-3.94900000	-2.85500000
H	1.34600000	-3.18100000	-0.57700000
H	2.99900000	-3.79600000	-0.80500000
H	1.16100000	3.63200000	-1.57700000
H	3.49100000	4.11000000	-1.99800000
H	3.49700000	3.13400000	-3.47800000
H	0.30000000	4.69700000	-3.53100000
C	-7.75700000	1.23700000	3.45300000

H	-7.71600000	2.31000000	3.22600000
H	-8.76100000	0.88900000	3.18300000
H	-7.64200000	1.11900000	4.53400000
C	-5.30500000	-5.25500000	3.09700000
H	-5.92700000	-5.34000000	2.19900000
H	-5.45200000	-6.17000000	3.68400000
H	-5.68300000	-4.41600000	3.68800000
C	-5.21500000	-6.15900000	-4.31400000
H	-5.06700000	-7.23400000	-4.15000000
H	-6.27300000	-5.94200000	-4.13500000
H	-5.00300000	-5.95700000	-5.36900000
C	-3.81800000	7.89400000	-1.53900000
H	-4.90200000	8.01100000	-1.65800000
H	-3.53800000	8.40100000	-0.60700000
H	-3.33300000	8.42400000	-2.36500000
C	-0.78000000	5.64200000	1.82100000
H	-1.38800000	5.46500000	0.92700000
H	-1.13900000	6.56700000	2.28400000
H	0.26700000	5.76500000	1.54500000
C	4.90300000	-1.57000000	3.06100000
H	4.47400000	-2.38200000	3.66500000
H	4.73200000	-1.82500000	2.02200000
H	5.97500000	-1.49800000	3.27000000
C	9.31400000	-1.86400000	2.37600000
H	9.95100000	-2.35000000	1.62900000
H	9.95200000	-1.35900000	3.10100000
H	8.72100000	-2.61700000	2.90400000
C	3.36000000	0.72800000	-4.58500000
H	4.21400000	1.18900000	-5.08400000
H	3.07300000	-0.19100000	-5.10100000
H	2.51000000	1.40700000	-4.54100000
O	2.85800000	3.45000000	3.04700000
H	1.91100000	3.68400000	3.14100000
H	2.95500000	3.25700000	2.07700000
O	2.83000000	3.08800000	0.38000000
H	2.54900000	3.99000000	0.17100000
H	1.84800000	2.17300000	0.22200000
N	1.13600000	1.22400000	0.14300000
C	1.04000000	0.77300000	1.47900000
C	-0.30900000	0.29200000	2.01300000
C	-0.42100000	-1.19700000	1.61400000
H	-1.41700000	-1.56400000	1.86400000
H	-0.26000000	-1.33900000	0.54300000
H	0.34000000	-1.77400000	2.14300000
C	-0.32500000	0.43100000	3.54500000
H	-0.23500000	1.48100000	3.83800000
H	-1.26800000	0.03100000	3.93300000
H	0.50100000	-0.12200000	4.00100000
C	-1.47700000	1.08200000	1.40100000
H	-2.40600000	0.81000000	1.90800000
H	-1.34100000	2.15900000	1.51600000
H	-1.61800000	0.84200000	0.34600000
H	0.21500000	1.45700000	-0.22000000

Table S17. Cartesian coordinates for EP FeNH

C	-2.83400000	-4.40300000	-1.82000000
O	-1.61100000	-4.19900000	-1.93900000
N	-3.74100000	-3.87100000	-2.66500000
S	3.71000000	-2.30200000	2.45700000

N	4.29300000	0.71300000	3.74300000
C	3.30600000	1.58800000	3.12600000
C	3.54200000	1.98900000	1.65600000
S	3.49200000	0.58300000	0.48100000
C	1.87300000	1.06900000	3.37700000
O	1.23300000	1.58200000	4.29900000
O	2.73200000	1.15400000	-0.71900000
O	4.96900000	0.30000000	0.11300000
N	1.34000000	0.09900000	2.56100000
C	-0.05800000	-0.23900000	2.88700000
C	-0.43700000	-1.52600000	2.19000000
O	-1.57900000	-2.01300000	2.25800000
C	-1.06200000	0.86500000	2.49800000
O	-0.91100000	1.25900000	1.12500000
N	0.57000000	-2.08900000	1.48900000
C	0.24300000	-3.34200000	0.80100000
C	1.53700000	-3.95400000	0.27000000
S	2.68400000	-2.66300000	-0.33500000
C	-0.41100000	-4.34400000	1.77300000
O	0.16200000	-4.73000000	2.77100000
O	0.63800000	-0.52800000	-3.76800000
C	-5.91800000	-4.75000000	2.43300000
C	-4.75500000	-4.30000000	3.01900000
C	-6.10600000	-3.94100000	1.25500000
N	-4.21500000	-3.25500000	2.28700000
C	-5.02800000	-3.00600000	1.20600000
C	-7.06800000	-3.92400000	0.22900000
C	-4.91100000	-2.05500000	0.18700000
C	-6.94800000	-2.98800000	-0.79800000
C	-5.88500000	-2.05500000	-0.81000000
C	-6.36300000	2.04000000	-1.70500000
C	-5.54200000	2.97100000	-2.36000000
C	-5.84900000	1.41800000	-0.55900000
C	-4.25900000	3.25700000	-1.90400000
C	-4.56100000	1.69400000	-0.08800000
C	-3.75300000	2.60900000	-0.77200000
O	-2.46300000	2.88700000	-0.40700000
N	3.75000000	-0.60000000	-3.24000000
C	3.77200000	0.73100000	-3.31000000
N	4.83900000	1.41800000	-2.81400000
N	2.78500000	1.38000000	-3.93700000
C	-1.13800000	6.38800000	3.98100000
C	-0.04000000	5.56500000	4.27800000
C	-1.75100000	6.22700000	2.72900000
C	0.46500000	4.65200000	3.35000000
C	-1.27000000	5.30800000	1.79000000
C	-0.13900000	4.55000000	2.09800000
O	0.41400000	3.65900000	1.17000000
C	-1.32400000	5.72600000	-2.33300000
C	-1.07200000	4.63500000	-3.17700000
C	-0.31100000	6.08800000	-1.42900000
C	0.11600000	3.90300000	-3.09700000
C	0.89100000	5.37500000	-1.34900000
C	1.09700000	4.26200000	-2.17300000
O	2.30100000	3.55600000	-2.12700000
N	7.97300000	-3.09400000	-1.12300000
C	7.36200000	-2.00900000	-0.61400000
N	7.65700000	-0.78400000	-1.07800000
N	6.45300000	-2.10900000	0.36000000
Fe	2.16700000	-1.13400000	1.36800000

H	-4.72200000	-3.86400000	-2.41600000
H	-3.43000000	-3.19600000	-3.34900000
H	-2.07100000	0.46100000	2.61800000
H	-0.93000000	1.72800000	3.15300000
H	-0.25900000	1.99000000	1.12000000
H	-0.17100000	-0.39900000	3.96800000
H	-4.26900000	-4.64200000	3.92200000
H	-3.34200000	-2.77300000	2.48200000
H	-7.89800000	-4.62600000	0.24100000
H	-7.69600000	-2.95300000	-1.58600000
H	-4.09400000	-1.34200000	0.17700000
H	-5.83500000	-1.31300000	-1.60100000
H	-6.45600000	0.69300000	-0.02400000
H	-5.91100000	3.47600000	-3.25100000
H	-4.18400000	1.18900000	0.79700000
H	-3.62300000	3.95400000	-2.43200000
H	-2.14300000	2.26800000	0.28500000
H	2.82100000	-1.01700000	-3.26500000
H	4.74800000	2.42600000	-2.87600000
H	5.11500000	1.11300000	-1.88100000
H	2.57900000	2.31800000	-3.58900000
H	1.97500000	0.81000000	-4.19400000
H	-2.62600000	6.82100000	2.48000000
H	0.43500000	5.63700000	5.25300000
H	-1.77100000	5.16700000	0.83800000
H	1.29700000	4.00500000	3.60400000
H	0.19200000	3.97700000	0.27700000
H	-0.46000000	6.93800000	-0.76900000
H	-1.82100000	4.33400000	-3.90400000
H	1.67400000	5.68200000	-0.66200000
H	0.27400000	3.04900000	-3.74700000
H	2.31100000	2.87500000	-1.42100000
H	7.75000000	-3.98200000	-0.69600000
H	6.09100000	-3.00100000	0.66600000
H	5.87100000	-1.28600000	0.57400000
H	8.37500000	-0.62700000	-1.76700000
H	7.12900000	0.01000000	-0.73400000
H	5.20100000	1.16100000	3.81300000
H	4.38900000	-0.18000000	3.26400000
H	3.32400000	2.52800000	3.68500000
H	2.75600000	2.67000000	1.31300000
H	4.52800000	2.44300000	1.52200000
H	-0.45300000	-3.14500000	-0.02100000
H	1.36100000	-4.66500000	-0.54200000
H	2.09100000	-4.44300000	1.07600000
H	-1.40900000	-4.72200000	1.49700000
C	-7.75500000	1.74900000	-2.21500000
H	-8.11800000	0.78000000	-1.85600000
H	-8.47000000	2.51000000	-1.87900000
H	-7.78500000	1.73600000	-3.31000000
C	-2.65000000	6.46100000	-2.36400000
H	-3.29700000	6.12500000	-1.54500000
H	-2.50800000	7.54100000	-2.25100000
H	-3.18400000	6.28600000	-3.30300000
C	-1.62200000	7.43100000	4.96700000
H	-1.06400000	8.36900000	4.84800000
H	-2.68300000	7.65900000	4.82100000
H	-1.48600000	7.09700000	6.00000000
C	-6.83500000	-5.82800000	2.93200000
H	-7.83100000	-5.42800000	3.16100000

H	-6.96900000	-6.62200000	2.18700000
H	-6.44300000	-6.28900000	3.84400000
C	-3.37000000	-5.26900000	-0.70900000
H	-4.43700000	-5.48000000	-0.80000000
H	-2.80500000	-6.20400000	-0.69000000
H	-3.21400000	-4.75400000	0.24200000
C	4.82200000	-1.40400000	-2.66500000
H	4.87000000	-1.29000000	-1.58300000
H	4.62400000	-2.44800000	-2.90900000
H	5.77600000	-1.10600000	-3.10600000
C	8.93900000	-3.10000000	-2.23500000
H	9.86000000	-2.57300000	-1.96400000
H	9.19000000	-4.13700000	-2.45400000
H	8.50000000	-2.65400000	-3.13300000
C	2.84600000	-2.69800000	4.02900000
H	3.53300000	-3.27300000	4.65400000
H	2.58000000	-1.76800000	4.53400000
H	1.95400000	-3.29000000	3.81300000
O	0.58600000	-3.46600000	-3.31900000
H	-0.29400000	-3.73200000	-2.95600000
H	0.43300000	-2.69900000	-3.89700000
O	1.88400000	-2.06400000	-1.65900000
H	1.40700000	-2.76000000	-2.24100000
H	1.13900000	0.39000000	-1.47600000
N	0.15700000	0.35700000	-1.74100000
C	-0.21200000	-0.15500000	-2.92000000
C	-1.71800000	-0.20200000	-3.23200000
C	-2.14200000	1.23100000	-3.63000000
H	-3.21500000	1.25400000	-3.83900000
H	-1.94400000	1.94000000	-2.82500000
H	-1.60800000	1.55800000	-4.53000000
C	-1.97800000	-1.15000000	-4.41300000
H	-1.71500000	-2.18400000	-4.16600000
H	-3.04100000	-1.11800000	-4.67800000
H	-1.39700000	-0.86200000	-5.29300000
C	-2.52300000	-0.66200000	-2.00000000
H	-3.58200000	-0.73400000	-2.25900000
H	-2.18500000	-1.64000000	-1.64600000
H	-2.46100000	0.04700000	-1.17100000
H	-0.51100000	0.60800000	-1.02300000

Table S18. Cartesian coordinates for E+P FeNH

C	3.95000000	-4.31600000	1.23600000
O	2.72400000	-4.14300000	1.28500000
N	4.74500000	-4.11200000	2.31900000
S	-5.57500000	-2.04700000	-1.23000000
N	-5.96700000	1.36000000	-0.56100000
C	-4.85900000	2.32300000	-0.62600000
C	-4.01400000	2.23400000	0.65500000
S	-3.84000000	0.48700000	1.20400000
C	-4.00000000	2.09500000	-1.88600000
O	-3.86900000	3.04600000	-2.67000000
O	-2.46200000	0.47100000	1.94800000
O	-4.98500000	0.31400000	2.21400000
N	-3.38000000	0.89200000	-2.05000000
C	-2.45800000	0.85000000	-3.20000000
C	-2.11300000	-0.58400000	-3.59500000
O	-1.39800000	-0.80700000	-4.58100000
C	-1.16000000	1.66100000	-2.97200000

O	-0.58000000	1.42400000	-1.66200000
N	-2.62400000	-1.55300000	-2.79700000
C	-2.18500000	-2.90200000	-3.14200000
C	-2.83300000	-3.93900000	-2.22000000
S	-2.76700000	-3.45000000	-0.46200000
C	-2.56900000	-3.24200000	-4.59800000
O	-3.72700000	-3.27000000	-4.96000000
O	0.69100000	-0.10300000	1.62500000
C	7.50800000	-3.40500000	-2.74400000
C	6.87100000	-2.18700000	-2.72400000
C	7.83300000	-3.70800000	-1.37000000
N	6.77800000	-1.71300000	-1.42000000
C	7.37500000	-2.62000000	-0.57100000
C	8.45900000	-4.79800000	-0.73900000
C	7.55300000	-2.58400000	0.81600000
C	8.62200000	-4.77800000	0.64400000
C	8.18100000	-3.67700000	1.41400000
C	6.99500000	2.60600000	1.01800000
C	5.77600000	3.29300000	1.13000000
C	7.23500000	1.90000000	-0.16900000
C	4.81900000	3.26100000	0.11600000
C	6.28800000	1.85900000	-1.20000000
C	5.06900000	2.52800000	-1.04800000
O	4.05200000	2.46700000	-1.97200000
N	-2.41800000	-2.22600000	2.76500000
C	-1.39300000	-3.07200000	2.64100000
N	-1.61800000	-4.41200000	2.70200000
N	-0.13900000	-2.62800000	2.53200000
C	0.05700000	6.96400000	-2.77500000
C	-1.15600000	6.27100000	-2.92100000
C	0.90400000	6.57900000	-1.72700000
C	-1.51400000	5.22900000	-2.06600000
C	0.56900000	5.52800000	-0.86400000
C	-0.63900000	4.85500000	-1.04600000
O	-1.00400000	3.76700000	-0.25000000
C	1.15900000	5.44000000	3.08500000
C	1.24400000	4.13800000	2.58300000
C	-0.08400000	5.84400000	3.61000000
C	0.15000000	3.25400000	2.60800000
C	-1.18000000	4.98900000	3.64400000
C	-1.07200000	3.67600000	3.15400000
O	-2.17500000	2.88900000	3.23600000
N	-7.35200000	-3.71000000	1.85600000
C	-7.12200000	-2.48300000	2.35900000
N	-7.07600000	-2.31200000	3.70000000
N	-6.93000000	-1.43700000	1.56500000
Fe	-3.52400000	-0.89300000	-1.16000000
H	5.75200000	-4.15800000	2.24600000
H	4.33000000	-3.77300000	3.17500000
H	-0.42800000	1.36000000	-3.72400000
H	-1.37200000	2.72400000	-3.07100000
H	-0.78600000	2.20700000	-1.10500000
H	-2.93700000	1.30400000	-4.07700000
H	6.47200000	-1.61200000	-3.54800000
H	6.40200000	-0.81900000	-1.14200000
H	8.81300000	-5.64200000	-1.32300000
H	9.11000000	-5.61100000	1.14100000
H	7.22000000	-1.73300000	1.40300000
H	8.35900000	-3.67100000	2.48600000
H	8.17700000	1.37400000	-0.30000000

H	5.56700000	3.86100000	2.03300000
H	6.50700000	1.32500000	-2.12300000
H	3.87700000	3.79300000	0.20900000
H	4.36600000	2.09800000	-2.80900000
H	-2.25200000	-1.22000000	2.61500000
H	-0.85100000	-5.01000000	2.42700000
H	-2.52000000	-4.74200000	2.38900000
H	0.57700000	-3.25400000	2.17200000
H	0.06900000	-1.63200000	2.38600000
H	1.84700000	7.09900000	-1.58000000
H	-1.84000000	6.55200000	-3.71800000
H	1.23800000	5.24000000	-0.05800000
H	-2.45600000	4.70400000	-2.18900000
H	-0.53800000	3.79800000	0.60700000
H	-0.19200000	6.85300000	4.00200000
H	2.18300000	3.78900000	2.16300000
H	-2.13100000	5.31300000	4.05400000
H	0.25400000	2.23100000	2.25400000
H	-2.06200000	1.99800000	2.83500000
H	-7.24800000	-3.80300000	0.85200000
H	-6.79200000	-1.60400000	0.55700000
H	-6.31700000	-0.66000000	1.89400000
H	-7.32600000	-3.04900000	4.33900000
H	-7.04100000	-1.37200000	4.06800000
H	-6.55800000	1.43800000	-1.38500000
H	-6.55000000	1.53900000	0.25300000
H	-5.19200000	3.36500000	-0.72000000
H	-3.01600000	2.64800000	0.50500000
H	-4.48700000	2.73300000	1.50600000
H	-1.09100000	-2.96000000	-3.05900000
H	-2.35300000	-4.91600000	-2.32100000
H	-3.90100000	-4.03200000	-2.43600000
H	-1.73900000	-3.48700000	-5.28500000
C	8.01300000	2.65800000	2.13700000
H	8.83000000	1.95100000	1.97000000
H	8.45200000	3.65900000	2.22300000
H	7.55200000	2.42200000	3.10300000
C	2.34500000	6.38500000	3.09900000
H	2.10600000	7.33300000	2.60400000
H	2.65100000	6.62000000	4.12500000
H	3.20900000	5.94800000	2.58600000
C	0.41900000	8.08600000	-3.72500000
H	-0.32900000	8.88700000	-3.69600000
H	1.39000000	8.52500000	-3.47600000
H	0.46900000	7.72600000	-4.76000000
C	7.83200000	-4.24900000	-3.94700000
H	8.91400000	-4.39500000	-4.04700000
H	7.37400000	-5.24300000	-3.87400000
H	7.47000000	-3.78000000	-4.86700000
C	4.67300000	-4.75100000	-0.02800000
H	5.68400000	-5.11800000	0.15900000
H	4.07800000	-5.52100000	-0.52400000
H	4.74400000	-3.89400000	-0.70500000
C	-3.81000000	-2.61000000	3.02200000
H	-3.87600000	-3.26900000	3.89400000
H	-4.36100000	-1.69200000	3.21600000
H	-4.25700000	-3.10300000	2.14900000
C	-7.62900000	-4.91000000	2.64900000
H	-8.53500000	-4.78400000	3.25200000
H	-7.79800000	-5.73900000	1.96200000

H	-6.78400000	-5.16700000	3.29700000
C	-6.10600000	-1.54500000	-2.92400000
H	-6.07100000	-0.45800000	-3.01000000
H	-5.43400000	-1.99400000	-3.65800000
H	-7.12700000	-1.89600000	-3.08900000
O	-1.11100000	-3.33000000	-0.22400000
H	-0.91600000	-2.36500000	-0.31800000
O	-1.46300000	-0.65700000	-0.30900000
H	-1.06100000	0.06300000	-0.87800000
H	-1.44100000	-0.27300000	0.59900000
N	2.01100000	0.86500000	0.07900000
C	1.83600000	0.21500000	1.24800000
C	3.07900000	-0.04600000	2.12600000
C	3.35700000	1.24900000	2.92400000
H	4.21600000	1.09200000	3.58600000
H	3.59700000	2.08000000	2.26000000
H	2.49600000	1.52700000	3.53800000
C	2.77900000	-1.18300000	3.11700000
H	2.60500000	-2.12700000	2.59500000
H	3.64100000	-1.31200000	3.78100000
H	1.91100000	-0.95100000	3.73700000
C	4.31300000	-0.42200000	1.27900000
H	5.13700000	-0.69200000	1.94700000
H	4.10600000	-1.28400000	0.63600000
H	4.66900000	0.40800000	0.66200000
H	1.21500000	1.04800000	-0.52500000
H	2.91500000	1.15900000	-0.25700000

Paper III

"Why nature prefers OH-proline to proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations"

Luigi Sgrizzi, Mario Prejanò, Isabella Romeo,Tiziana Marino,
Nino Russo.

Manuscript submitted

Chemistry - A European Journal

Why nature prefers hydroxy-proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations

--Manuscript Draft--

Manuscript Number:	
Article Type:	Full Paper
Corresponding Author:	Tiziana Marino Univ Calabria Arcavacata Di Rende, CS ITALY
Corresponding Author E-Mail:	tmarino@unical.it
Order of Authors (with Contributor Roles):	Luigi Sgrizzi
	Mario Prejanò
	Tiziana Marino
	Isabella Romeo
	Nino Russo
Keywords:	polysaccharide deacetylase; zinc; density functional calculations; catalytic mechanism; hydroxy proline
Manuscript Classifications:	Computational chemistry; Density functional calculations; Enzyme catalysis
Suggested Reviewers:	Fahmi Himo fahmi.himo@su.se
	Pedro Alexandrino Fernandes pafernan@fc.up.pt
	Tomasz Borowski ncborows@cyf-kr.edu.pl
	Juan Manuel Bort andres@qfa.uji.es
	Max Holthausen max.holthausen@chemie.uni-frankfurt.de
Opposed Reviewers:	
Abstract:	Nature exploits different strategies for enhancing the catalytic activity of enzymes resorting to beneficial mutations. The case of the proline hydroxylation in active site of polysaccharide deacetylase from Bacillus Cereus Bc1960 is an interesting example as small chemical modifications causes significant improvements in enzymatic activities. Starting from preliminary Molecular Dynamics simulations on the enzyme-substrate complex we have undertaken QM cluster investigation to explore the deacetylation mechanism considering both OH-proline and proline amino acids in the enzyme active site. Our calculations highlight as the hydrogen bond network established by the –OH group on the Cα of the proline with its closer neighbours stabilizes the transition states and, consequently, the occurring of the reaction. The analysis of the obtained PESs reveals as also the intermediate and the product get energetic benefits from the presence of hydroxyl-proline. These results corroborate the experimental findings on the polysaccharide deacetylase members that propose a higher deacetylase activity by a factor of 10 with respect to the non-hydroxylated form.
Author Comments:	Arcavacata, November 30th, 2018 MANUSCRIPT TITLE: Why nature prefers hydroxy-proline in the deacetylation process promoted by peptidoglycan N-deacetylase: insight from molecular simulations AUTHORS: Luigi Sgrizzi, Mario Prejanò, Isabella Romeo, Tiziana Marino* and Nino Russo

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	MANUSCRIPT TYPE: Article To the Editorial Board of Chemistry, A European Journal Enclosed, please, find the above mentioned manuscript that we would like to publish in Chemistry A European Journal. The present manuscript has been submitted exclusively to Chemistry, A European Journal and all authors of the manuscript are aware of and have agreed to the submission decision. COMMENTS ON THE MANUSCRIPT SIGNIFICANCE Post-translational modifications of proteins can be considered an alternative way to increase the functional complexity of the proteome. Lately in the active sites of the putative polysaccharide deacetylases (PDA) members of the carbohydrate esterase family 4 involved in various mechanism of bacterial pathogens, a new type of modification involving the proline backbone has been observed. In particular, a hydroxyproline has been identified at the position of the Pro171 in a known PDA Bacillus cereus Bc1960 (Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 2008, 64,203) Following the most recent experimental evidences (J. Am. Chem. Soc. 2017, 139, 5330), we have undertaken a theoretical study devoted to provide a detailed atomistic understanding of the different catalytic activities observed between the hydroxyl-proline and proline containing enzyme. In the present work, the energy profiles computed by using first-principle density functional theory and applying the quantum chemical cluster methodology are calculated. In particular, attention was focused on the hydroxyl group of modified proline during the reaction. Corresponding Author Contact Information: -Tiziana Marino Dipartimento di Chimica e Tecnologie Chimiche, Ponte P. Bucci Cubo 14C, Università degli Studi della Calabria, 87036 Arcavacata di Rende (CS), Italy. Email: tmarino@unical.it Phone: +39-0984-492085 Fax: +39-0984-492044
Section/Category:	
Additional Information:	
Question	**Response**
Submitted solely to this journal?	Yes
Has there been a previous version?	No
Do you or any of your co-authors have a conflict of interest to declare?	No. The authors declare no conflict of interest.
Animal/tissue experiments?	No

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Click here to access/download;Manuscript;Manuscript.docx

Why nature prefers hydroxy-proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations

Luigi Sgrizzi, Mario Prejanò, Isabella Romeo, Nino Russo and Tiziana Marino[a]*

Abstract: Nature exploits different strategies for enhancing the catalytic activity of enzymes resorting to beneficial mutations. The case of the proline hydroxylation in active site of polysaccharide deacetylase from *Bacillus Cereus Bc1960* is an interesting example as small chemical modifications causes significant improvements in enzymatic activities. Starting from preliminary Molecular Dynamics simulations on the enzyme-substrate complex we have undertaken QM cluster investigation to explore the deacetylation mechanism considering both OH-proline and proline amino acids in the enzyme active site. Our calculations highlight as the hydrogen bond network established by the –OH group on the Cα of the proline with its closer neighbours stabilizes the transition states and, consequently, the occurring of the reaction. The analysis of the obtained PESs reveals as also the intermediate and the product get energetic benefits from the presence of hydroxyl-proline. These results corroborate the experimental findings on the polysaccharide deacetylase members that propose a higher deacetylase activity by a factor of 10 with respect to the non-hydroxylated form.

Introduction

Proteins cover a wide range of functions by using the same twenty amino acids combined in different ways. Since their biosynthesis, proteins can suffer covalent modifications by introduction of new functional groups not present in the standard amino acid building blocks making it happen a new chemical properties. Post-translational modifications (PTM) play an essential role in control mechanisms, stability and properties, going to modify protein function.[1] In the case of enzymes, when the modification occurs on the protein backbone, an improvement of the existing functions occurs with a consequent increase of their stability and activity. An interesting case of backbone modifications is the Cα hydroxylation of a proline in the active site of a *Bacillus cereus Bc*1960 peptidoglycan N-acetylglucosamine (GlcNAc) deacetylase, a known polysaccharide deacetylase (PDA) has been recently reported by *Fadouloglou et al.* [1]Evidences from cell-free expression performed in the absence of PDA substrate suggest that the Cα-hydroxylation of Pro to give 2-hydroxyproline (2-Hyp) is an autocatalytic hydroxylation process, where molecular oxygen is the source of the –OH group instead of waters. Other examples of autocatalytic hydroxylation exist in many other enzymes, where just side chains suffered modification. [2] 2-Hyp is also involved in plant cell wall architecture,[3] in signalling processes linked to hypoxia response,[4] and in physiological pathways combined with diseases such as cancer. [5] Furthermore, this kind of post-translational modifications represents an atypical modification process of PDA components because it establishes an intertwining between the hydroxylation and deacetylation reactions due to a deacetylation's enhancement by approximately a factor of 10 relative to the activity of the non-hydroxylated form.Peptidoglycan, the peptide-linked heteropolymer of GlcNAc and N-acetylmuramic acid is a protective component of the bacterial cell wall, [6] that plays a vital role in the maintenance of the stability and viability of bacterial cells. Different kinds of peptidoglycan modification, such as de-N-acetylglucosamine acetylationare involved in different biological functions such as bacterial growth, division and autolysis.[7] Such peptidoglycan modifications can be seen as a means to counteract the activity of the host lysozyme, which suggests a probable biological role of peptidoglycan deacetylases in the protection of bacteria from the defence mechanisms of the hosts. [7] As a consequence, the modified peptidoglycan is no further recognized as target by these enzymes.[8-10] PDAs are members of the carbohydrate esterase family 4 (CE4) and they play a decisive role in diverse mechanisms of bacterial pathogens.[11] The crystal structure of the *Bc1960* peptidoglycan N-acetylglucosamine deacetylase has been recently determined in the metal-free form at 2.3 Å resolution. [12] It presents in the active site a divalent Zn metal coordinated with the broadly conserved triad in the CE-4 family, [1,6,11,13] consisting of two His residues and one Asp (His131, His135 and Asp81 in *Bc1960*). The presence of a Pro residue in the vicinity of the metal (Pro171) with a Cα atom-metal distance of less than 8 Å and of Asp80, acting as catalytic residue, completes the coordination shell of the metal ion. As member of CE-4 family *Bc1960* PDA uses acid/basis catalysis His in the reaction. The presence of a hydroxyl proline could be linked to various aspects of bacterial pathogenicity and the catalytic mechanism of this kind of enzymes is largely unknown, and deeper insights on the mechanism of *Bc1960* PDA are highly desired. In fact, understanding the different properties (energetic and electronic) due to the presence of 2-Hyp with respect to the standard Pro can help to further elucidate the sources of discrimination between the two and to design mechanism-based inhibitors for this important class of enzymes. With this in mind, we have undertaken a QM cluster model study on the reaction mechanism considering both the free and hydroxylated prolines in the *Bc1960* PDA active site obtaining other than the relative reaction mechanism profile atomistic details that explain the improvement of enzymatic activity in the presence of the 2-Hyp.

Methods

Quantum chemical active site model The cluster model of the active site of the best selected enzyme-substrate complex (Figure 1) includes the whole coordination shell of the zinc ion

[a] L. Sgrizzi, M. Prejanò, Dr. I. Romeo, Prof. N. Russo, Prof T. Marino
Dipartimento di Chimica e Tecnologie Chimiche
Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)
E-mail: tiziana.marino65@unical.it

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

WILEY-VCH

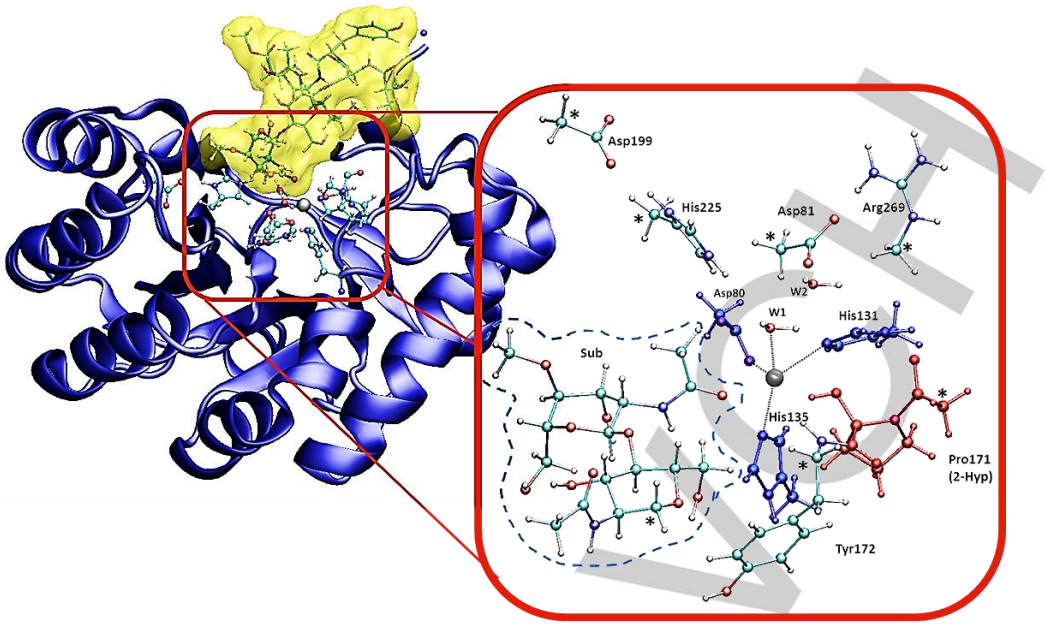


Figure 1. On the left, ES complex structure isolated from MD simulation of PDA-2-Hyp and on the right the arised cluster model adopted in QM investigation. C atoms labelled with "*" were kept frozen during the optimizations.

(His131, His135 and Asp80 and the water molecule, **w1**) and the binding cavity (Asp81, Asp199, Pro171 (**2-Hyp**), Tyr172 and Arg269). Another crystallographic water molecule was also included in the model (**w2** in Figure 1).

This water molecule does not participate in the chemical transformation but it is within hydrogen bonding distance of 2-Hyp and carbonyl of the substrate. To preserve the integrity of the active site model, 8 atoms were kept fixed to the starting positions during geometry optimizations. The substrate considered in the QM mechanistic study has been downsized with respect to that used in the MDs performed on the complex substrate-enzyme, retaining the N-acetyl-glucosamine moiety as shown in the Supplementary Material. The resulting cluster model consists of 178 atoms with a charge of +1.

Quantum Chemical Details All the QM calculations were performed with the Gaussian 09 code. [14] The B3LYP-D3 hybrid DFT functional [15,16] combined with the 6-31+G(d,p) basis set for all atoms and SDD for the zinc atom [17] was used for the optimizations of all the stationary points intercepted along the potential energy surface. The effect of protein environment, beyond the cluster model, was approximated by single-point energy calculation using the SMD solvation model [18] with a dielectric constant of ε = 4, in line with previous studies. [19-24] Zero-point energies (ZPE) were obtained by analytical frequency calculations on the optimized geometries. Both the ZPE and solvation calculations were performed at the same level of theory as the geometry optimization. To improve the energies, single-point calculations were performed with the larger 6-311+G(2d,2p) basis set. Thus, the reported B3LYP-D3 [25] energetic values are the energies calculated with the large basis set including both the ZPE and SMD solvation corrections.

Furthermore, the QM energies of all stationary points were refined subsequently by single point calculations (6-311+G(2d,2p) basis set for all atoms and SDD for Zn) with other three hybrid functionals (DFs) (M06,[26] M06-2x,[26] wB97X-D[27]) and taking into account the effect of protein (ε= 4) by SMD model on the optimized geometries at B3LYP-D3 level.

Results and Discussion

Since the employed metal free crystallographic structure of the alone of *Bc1960* enzyme (PDB ID: 4L1G)[11] did not contain substrate molecule, we believed necessary to perform MD simulations on the enzyme and then on the enzyme-substrate complex obtained by the best pose of docking procedure. This strategy allowed us to observe a different conformational behavior between 2-Hyp and Pro forms of the PDA enzyme. In fact, in the case of proline containing enzyme differently from 2-Hyp one, the substrate moves away from the catalytic site after 20 ns of MDs making difficult the Michaelis complex formation (see Figure S1) On the other hand since the active site is located in an extended surface accessible to the solvent the OH group of the hydroxyl proline can be more easily docked. (see Figure S2). A deeper analysis of H-bond interactions (see Figure 2) between donor and acceptor in the enzyme active site show as, contrary to the proline, the 2-Hyp attracts more water molecules that better interact with the hydrophilic region of the substrate (the glucosamine). Cluster analysis performed on the structures on extracted along the MD simulation allow to obtain a reliable starting structure for the ES complex for the 2-Hyp containing enzyme.

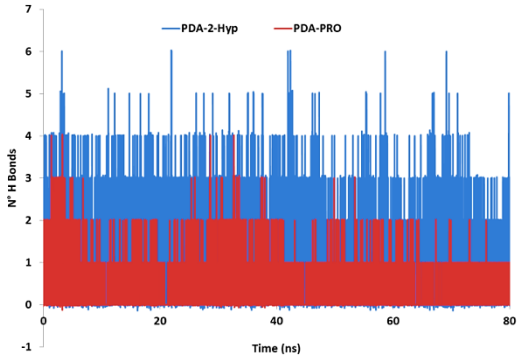
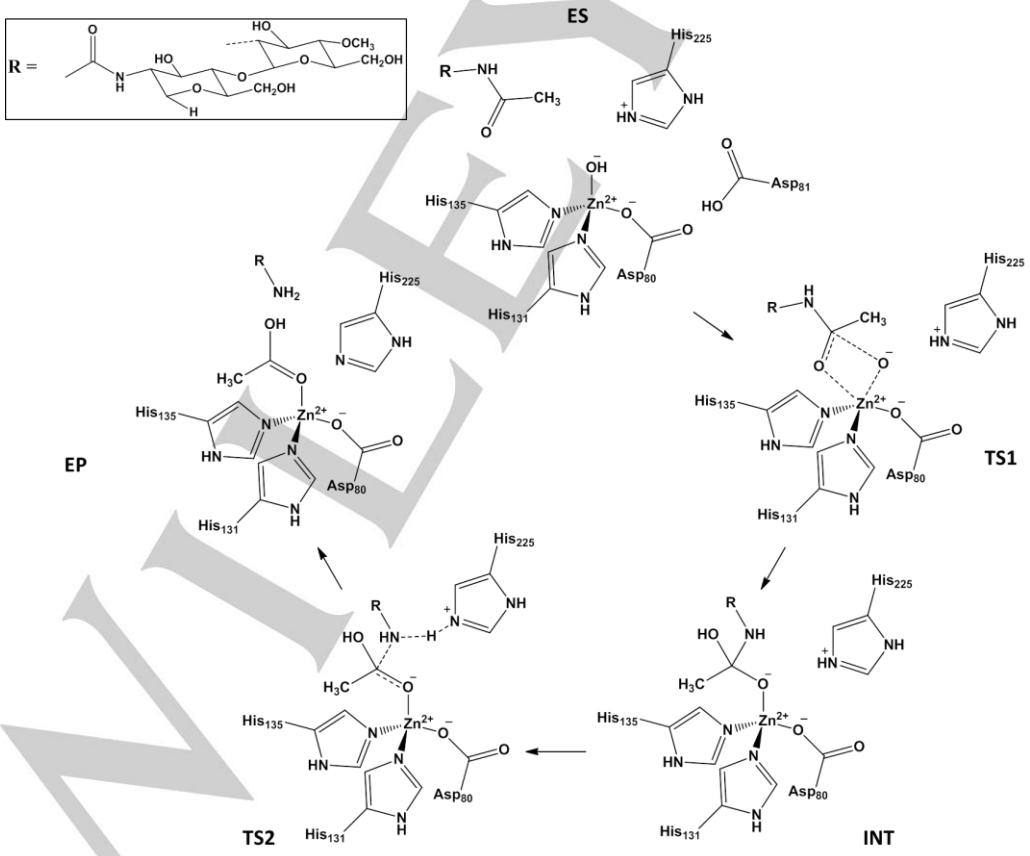


Figure 2. Hydrogen bonds number along MD simulation in the catalytic pocket of PDA-2-Hyp and PDA-PRO..

The proposed deacetylation mechanism of PDA *Bc1960* is based upon earlier work,[8] on zinc-dependent deacetylases: Zn^{2+} plays the role of Lewis acid in the activation of waters while Asp81 and His225 residues act as catalytic base and acid, respectively. Consequently, free amine and the acetate product are generated. The starting ES complex includes a water molecule **w1** coordinated to the zinc ion that transfers its proton to the Asp81. The proposed mechanism (Scheme 1) provides that the reaction occurs in two steps: in the first one the nucleophilic attack takes place (TS1) with the formation of the tetrahedral intermediate (INT) while in the last one (TS2) the C-N bond cleavage occurs (TS2) and the acetyl group coordinated to the metal ion. The obtained reaction energy profile for both 2-Hyp and Pro forms of the is reported in Figure 3.

In particular, the energetic profile with regards to the proline containing enzyme, the energetic profile has been obtained by single points on all the stationary points of the 2-Hyp enzyme. Although our MD simulation shows as in PDA-PRO the substrate possesses a greater mobility in the active site, we have also considered its reaction path to better elucidate the role of the OH moiety on the Cα position. In particular, the energetic profile of the proline containing enzyme has been obtained by single points computation on all stationary points intercepted for the 2-Hyp enzyme. The optimized Michaelis complex of 2-Hyp enzyme (Figure 4) shows forms two hydrogen bonds between the –OH moiety of 2-Hyp (2.285 Å) and the water molecule (**w2**) (1.997 Å) retained in the active site model. **w2** is implicated also in hydrogen bond (1.659 Å) with hydroxyl group coordinated to the zinc ion. These species are parts of the oxyanion hole indicating their roles in the initial binding and positioning of the acetyl group of the substrate, while the NH moiety of the backbone of Tyr172, experimentally proposed as transition state stabilizer, [6] is found at 3.406 Å from the oxygen of the substrate.



Scheme 1. Proposed mechanism for the deacetylation of GlnNAc by PDA-2-Hyp.

WILEY-VCH

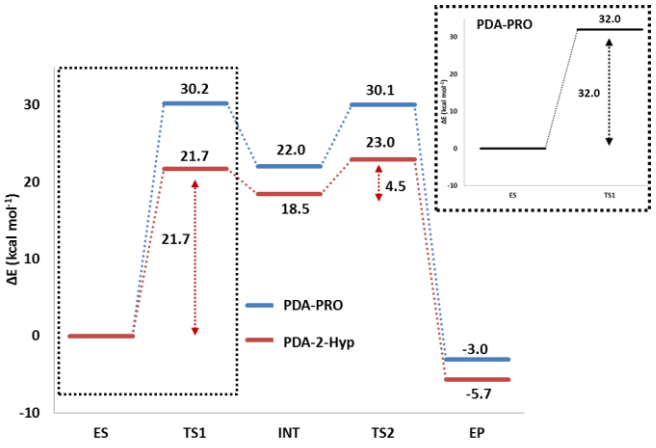


Figure 3. B3LYP-D3/6-311+G(2d,2p)!SDD//B3LYP-6-31+G(d,p) energies in protein environment for PDA-2-Hyp and PDA-PRO. In the box, energy barrier for ES-TS1 step for PDA-PRO is reported, in the case of relaxed geometries, at the same level of theory.

Furthermore, this carbonyl oxygen at 3.780 Å from the zinc ion oriented toward the metal center. The water hydroxyl group lies at a distance of 1.917 Å from the Zn and at 3.650 Å from the carbonyl substrate.

The first step of the deacetylation reaction involves OH nucleophilic attack on the substrate carbonyl carbon (TS1 of Figure 3 giving rise to negatively charged tetrahedral intermediate (INT of Figure 4). In TS1 the OH--C=O is now 1.651 Å, the hydrogen bond involving the 2-Hyp and oxygen carbonyl is now more reinforced since it is 2.102 Å and the forming acetic acid shows a bicoordinated fashion (2.177 Å and 2.083 Å for O and OH group, respectively) as occurred in other zinc containing enzymes. [23,24,28] The energy of TS1 is calculated to be 21.7 kcal/mol. INT1 lying at 18.5 kcal/mol above the ES, presents the already formed OH-C bond (1.484 Å) and the OH moiety of 2-Hyp is yet hydrogen bonded to the oxygen carbonyl (2.234 Å). The pre-existing hydrogen bonded water bridge between OH of 2-Hyp (1.917 Å) and the OH of acetyl group (1.739 Å) is retained.

The C-N scissile bond is now 1.462 Å longer than that in ES species (1.354 Å). Also in this case the –NH moiety from Tyr172 does not contribute to the H-bonds network above mentioned since its distance from the oxygen of the acetyl group is 4.397 Å. The next step describes the C-N cleavage for generating the acetic acid and the amine counterpart. This occurs *via* a concerted transition state (TS2) in which the C− N bond breaks (1.574 Å) and a proton is transferred from the protonated His225 to the nitrogen of the "leaving group". (see Figure 6) As consequence, the OH-Zn^{2+} distance undergoes an elongation (1.951 Å) and C-O a shortening (1.352 Å). The energy of TS2 is found to be 5.5 kcal/mol higher than INT species. In EP complex, the acetic acid product is still coordinated to the zinc by carbonyl moiety (1.971 Å) and hydrogen-bonded to 2-Hyp (1.760 Å) and the amino fragment generated by C-N cleavage is hold in catalytic site by hydrogen bond with the Asp80 (1.856 Å). w2 remains trapped in hydrogen bond bridge between the OH of 2-Hyp (2.069 Å) and that of acetic acid (1.895 Å). The NH of the backbone of Tyr172 is far from w2 3.501 Å. The energy of the enzyme−product complex lies at 5.7 kcal/mol below the ES (Figure 3).

Since the His225, active in the acid/base catalysis is oriented to the OH group of the acetic acid (2.176 Å), a proton must be transferred to this amino acid and the entering water molecule (w2) moves the acetate from the metallic site to restart the catalytic cycle.

According to these results, TS1 seems to be is

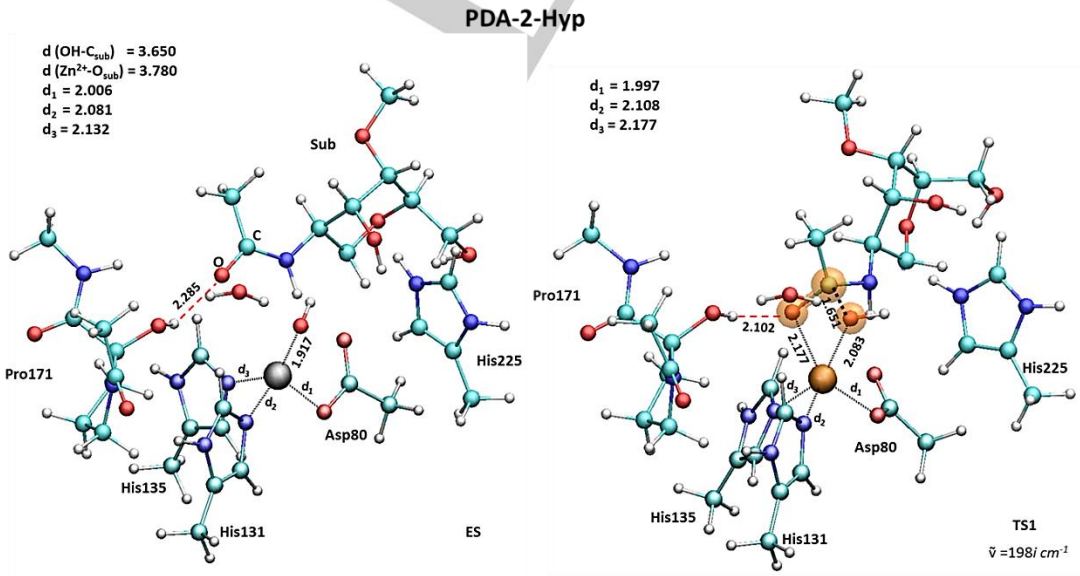


Figure 4. B3LYP/6-31+G(d,p) optimized structures of ES and TS1 stationary points of PDA-2-Hyp system.

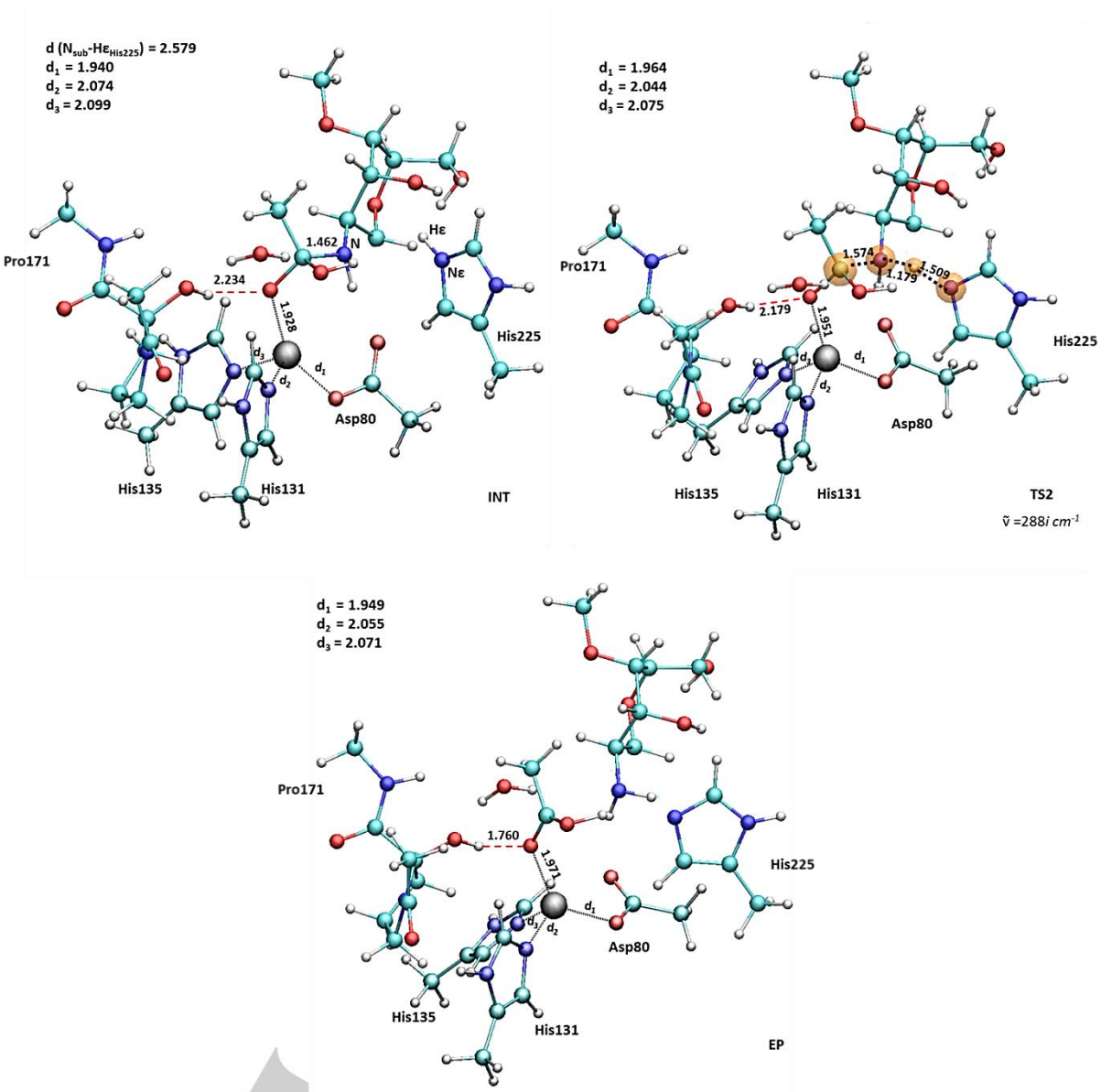


Figure 5. B3LYP/6-31+G(d,p) optimized structures of INT, TS2 and EP stationary points for the PDA-2-Hyp.

the rate-limiting step with a barrier of 21.7 kcal/mol, relative to the ES complex (Figure 3).

Going to discuss the PES for the proline containing enzyme, (Figure 3) it is possible to observe as the presence of proline rather than 2-Hyp influences the whole mechanism. In particular, an increase on energy by 8.5 kcal/mol and 7.1 kcal/mol for TS1 and TS2, respectively, occurred. Moreover, also the intermediate undergoes a higher destabilization with a consequent enhanced INT-TS2 gap (8.1 kcal/mol versus 4.5 kcal/mol present in the 2-Hyp PDA). The lack of the hydrogen bonds in all the intercepted accounts for the better catalytic activity of PDA-2-Hyp.

In order to verify if this result is depending by the use of frozen geometries used in the energy profile computation of proline containing enzyme, we redone the calculations relaxing the geometries from the PDA in the first step of the mechanism. The obtained optimized structures for ES and TS1 are shown in Figure 6. It can be noted that in ES the distances involving the nucleophilic agent and carbon carbonyl and the Zn^{2+} and the oxygen carbonyl of the substrate (4.376 Å and 4.661 Å, respectively) result to be longer than the corresponding ones in the 2-Hyp PDA. In addition, the lack between the NH group of Tyr172 of the hydrogen bond and the substrate carbonyl (2.438

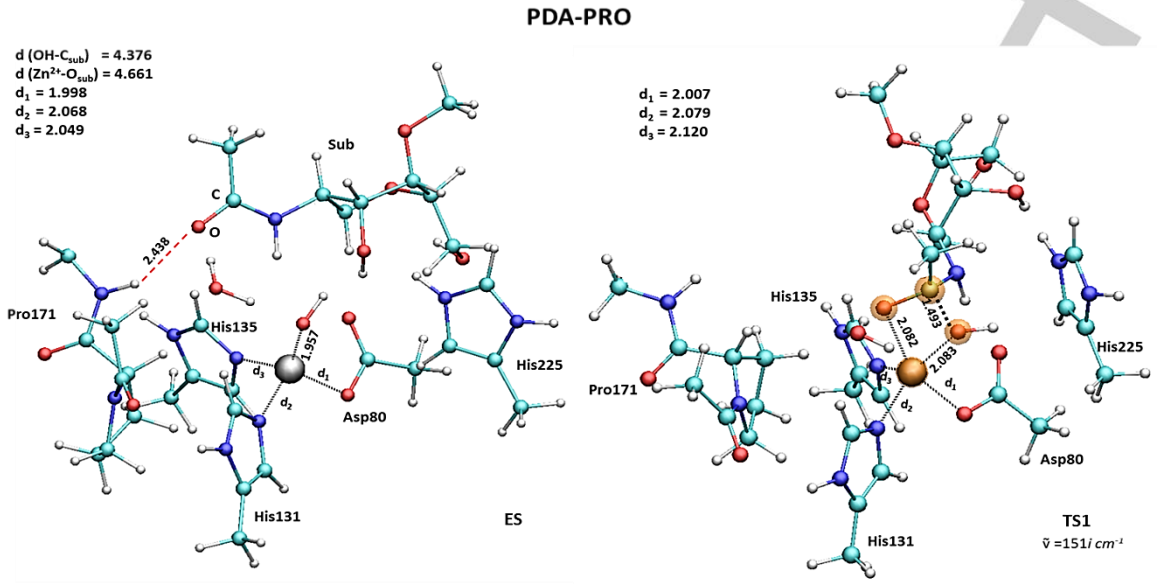


Figure 6. B3LYP/6-31+G(d,p) optimized structure of ES and TS1 stationary points of PDA-PRO system.

Å) destabilizes the transition state and the barrier becomes 32.0 kcal/mol.

In the case of the 2-Hyp containing enzyme, QM energies were refined by single point calculations with other three hybrid DFs as above reported in quantum chemical details. Results in comparison with the B3LYP-D3 ones are collected in Table S1 in the SI. All the density functionals showed a similar overall agreement for the energetic profile, although an improvement of the barriers has been obtained using the M06-2x functional.

Conclusions

This work was devoted to understand the catalytic mechanism of 2-Hyp-PDA at the atomistic level, and to investigate the effects induced by the presence of the post-translational modified proline having the hydroxyl group on the Cα backbone by examining the elementary steps of the catalytic process. In particular, in the present study we have investigated the deacetylation process promoted by PDA-2-Hyp enzyme performing a fully QM investigation performed in the framework of density functional theory and using the cluster approach. Preliminary MD simulations have been carried out for generating the enzyme-substrate complex be subjected to the quantum mechanical study. The explored mechanism takes into account the most recent experimental observations on this class of metalloenzymes that propose the hydroxyl group from w1 coordinated to the zinc ion as nucleophile agent during the reaction. Furthermore, our investigation suggests the importance of the modified proline along with a water molecule present in the active site as indirect assistants during the catalytic reaction.

Because of the outcomes of our study, it is possible to draw the following conclusions:

- The transition state describing the nucleophilic attack (TS1) represents the rate-limiting step.
- In every stationary point of the energetic profile, the OH group of modified proline is implicated in a hydrogen bond network with its closer neighbors. At this purpose, the stabilizing contribution of the hydrogen bonds by 2Hyp has been evaluated for comparison with that of proline containing PDA, arising from single point calculations on all the stationary points of the PDA-2-Hyp removing the OH group of 2-Hyp.
- The ES-TS1 relaxed calculated barrier of PDA-PRO (32.0 kcal/mol) evidences the lacking of stabilizing hydrogen bonds corroborating what suggested from single point energy calculations.
- Our results provide important atomistic details of the deacetylation mechanism of a polysaccharide deacetylase member useful for driving future development of mechanism-based inhibitors for this important class of enzymes.

Acknowledgements

Financial support from the Universita degli Studi della Calabria, Dipartimento di Chimica e Tecnologie Chimiche (CTC), is acknowledged.

Conflict of interest

The authors declare no competing financial interest.

Keywords: polysaccharide decetylase •Zinc • density functional calculations • catalytic mechanism•hydroxy proline

[1] V. E. Fadouloglou, S. Balomenou, M. Aivaliotis, D. Kotsifaki, S. Arnaouteli, A. Tomatsidou, G. Efstathiou, N. Kountourakis, S. Miliara, M. Griniezaki, A. Tsalafouta, S. A. Pergantis, I. G. Boneca, N. M. Glykos, V. Bouriotis, M. Kokkinidis, *J. Am. Chem. Soc.* **2017**, *139*, 5330-5337.

[2] A. Liu, R. Y. Ho, L. Jr. Que, M. J. Ryle, B. S. Phinney, R. P. Hausinger, *J. Am. Chem. Soc.* **2001**, *123*, 5126-5127.

[3] D. T. Lamport, M. J. Kieliszewski, Y. Chen, M. C. Cannon, *Plant Physiol.* **2011**, *156*, 11-19.

[4] B. F. Palmer, D. Clegg, *J. Mol. Cell. Endocrinol.* **2014**, *397*, 51-58.

[5] N. Masson, P. Ratcliffe, *J. Cancer Metab* **2014**, *2*, 3.

[6] D. E. Blair, A. W. Schüttelkopf, J. I. MacRae, D. M. F. van Aalten, *Proc. Natl Acad. Sci. USA* **2005**, *102*, 15429-15434.

[7] I. G. Boneca, O. Dussurget, D. Cabanes, M. A. Nahori, S. Sousa, M. Lecuit, E. Psylinakis, V. Bouriotis, J. P. Hugot, M. Giovannini, A. Coyle, J. Bertin, A. Namane, J. C. Rousselle, N. Cayet, M. C. Prévost, V. Balloy, M. Chignard, D. J. Philpott, P. Cossart, S. E. Girardin, *Proc. Natl Acad. Sci. USA* **2007**, *104*, 997–1002.

[8] I. G. Boneca, *Curr. Opin. Microbiol.* **2005**, *8*, 46–53.

[9] A. Severin, K. Tabei, A. Tomasz, *Microb. Drug Resist.-Mechan. Epidemiol. Dis.* **2004**, *10*, 77–82.

[10] A. Bera, S. Herbert, A. Jakob, W. Vollmer, F. Gotz, *Mol. Microbiol.* **2005**, *55*, 778–787.

[11] S. Arnaouteli, P. Giastas, A. Andreou, M. Tzanodaskalaki, C. Aldridge, S. J. Tzartos, W. Vollmer, E. Eliopoulos, V. Bouriotis, *J. Biol. Chem.* **2015**, *290*, 13465-13478.

[12] A. Tsalafouta, E. Psylinakis, E. G. Kapetaniou, D. Kotsifaki, A. Deli, A. Roidis, V. Bouriotis, M. Kokkinidis, *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **2008**, *64*, 203-205.

[13] V. E. Fadouloglou, M. Kapanidou, A. Agiomirgianaki, S. Arnaouteli, V. Bouriotis, N. M. Glykos, M. Kokkinidis, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69,* 276-283.

[14] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Jr. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2013.

[15] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.

[16] C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[17] D. Andrae, H. Haußermann, M. Dolg, H. Stoll, H. Preuß, *Theor. Chim. Acta* **1990**, *77*, 123–141.

[18] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378−6396.

[19] M. Prejanò, T. Marino, N. Russo, *Chem. Eur. J.* **2017,** *23*, 8652-8657.

[20] M. Prejanò, T. Marino, C. Rizzuto, J. C. Madrid Madrid, N. Russo, M. Toscano, *Inorg. Chem.* **2017**, 56, 13390-13400.

[21] O. Amata, T. Marino, N. Russo, M. Toscano, *Phys. Chem. Chem. Phys.* **2011**, *13*, 3468−3477.

[22] M. E. Alberto, T. Marino, N. Russo, E. Sicilia, M. Toscano, *Phys. Chem. Chem. Phys.* **2012**, *14*, 14943−14953.

[23] P. Piazzetta, T. Marino, N. Russo, D. R. Salahub, *ACS Catal.* **2015,** *5*, 5397-5409.

[24] P. Piazzetta, T. Marino, N. Russo, *Phys. Chem. Chem. Phys.* **2014**, 16, 16670−16676.

[25] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215-241.

[26] J. D. Chain, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615-6620.

[27] M. Hernick, C A. Fierke, *Arch. Biochem. Biophys.* **2005**, *433*, 71-84.

[28] P. Piazzetta, T. Marino, N. Russo, *Inorg. Chem.* **2014**, *53*, 3488-3493.

WILEY-VCH

WILEY-VCH

Click here to access/download
Supporting Information
ZN_SI.docx

Why nature prefers hydroxy-proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations

Luigi Sgrizzi, Mario Prejanò, Isabella Romeo, Tiziana Marino* and Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via P. Bucci, I-87036-Arcavacata di Rende, Italy

*Correspondence:

Tiziana Marino

tiziana.marino65@unical.it

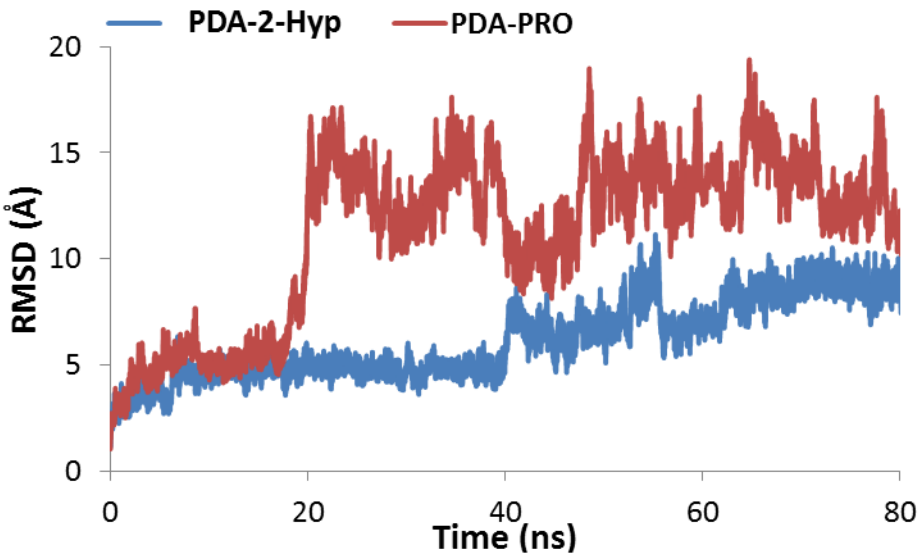


Figure S1. RMSD of substrate in ES complex MD of PDA-2-Hyp (blue line) and PDA-PRO (red line).

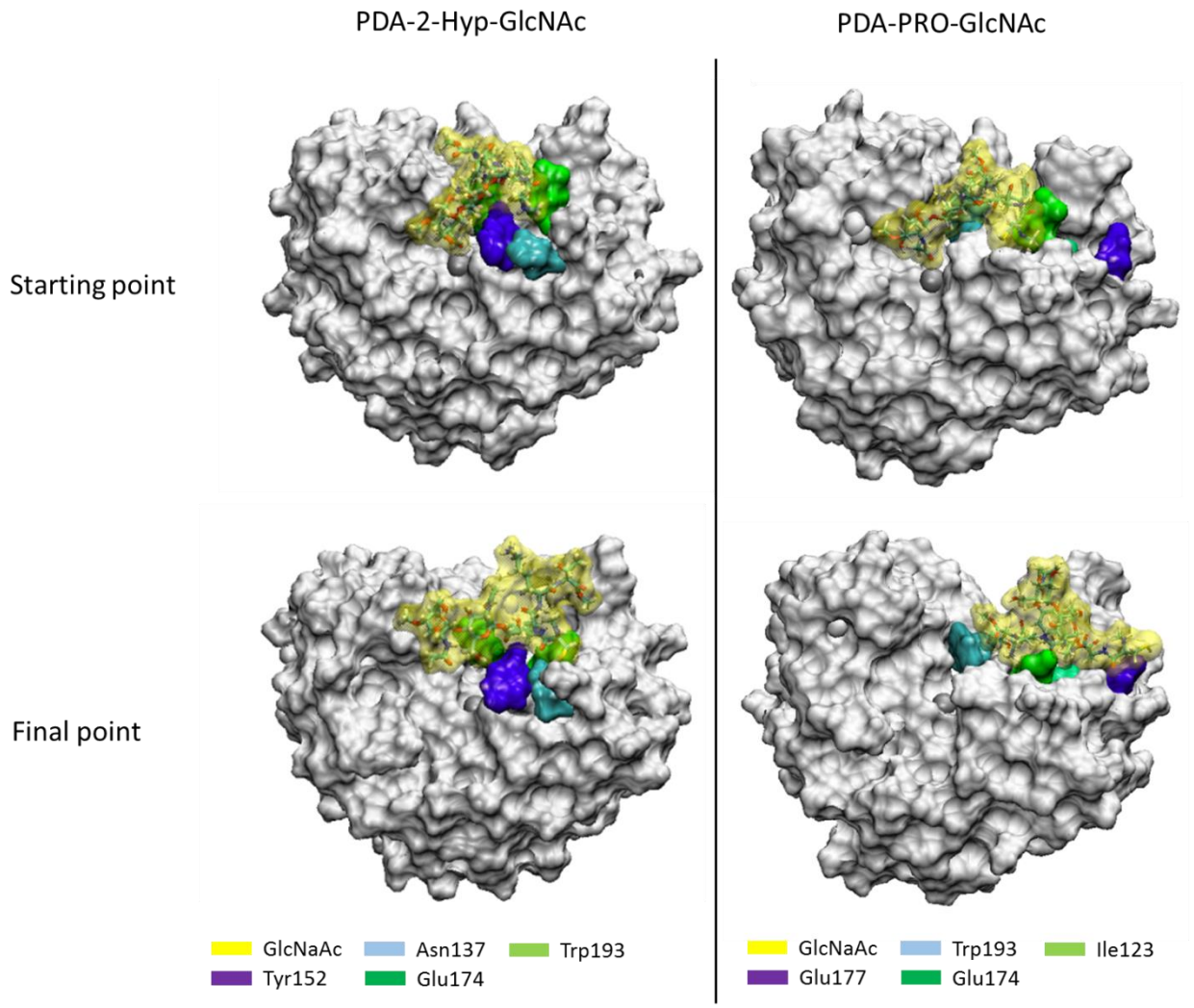


Figure S2. Starting and final points of 80 ns MD of PDA-2-Hyp-GlcNAc and PDA-PRO-Glc-NAc. The frequencies of hydrogen bonds is marked with different colors: 25 % (purple), 20% (cyan), 15 % (green) and 10 % (light green).

Parametrization of peptidoglycan N-acetylglucosamine

The peptidoglycan N-acetylglucosamine (GlcNAc) deacetylase BC1960 from *Bacillus cereus* (EC 3.5.1.33), is an enzyme consisting of 275 amino acids. [1] As other members of CE4 family it catalyzes the hydrolysis of N-linked acetyl groups from N-acetylglucosamine residues (chitin deacetylases, NodB chitooligosaccharide deacetylases and peptidoglycan N-acetylglucosamine deacetylases). [2]In order to perform MM and MD computations it is mandatory to determine the force field parameters for the peptidoglycan N-acetylglucosamine substrate (Figure S3). From optimized structures at the HF/6-31G(d) level of theory they have been obtained with the Antechamber tool, as implemented in Amber16[3] and by using the GAFF force field[4] to derive the intramolecular and Lennard-Jones parameters, and to derive atomic charges at the HF/6-31G(d) level of theory throughout the RESP (restrained electrostatic potential) method. [5]

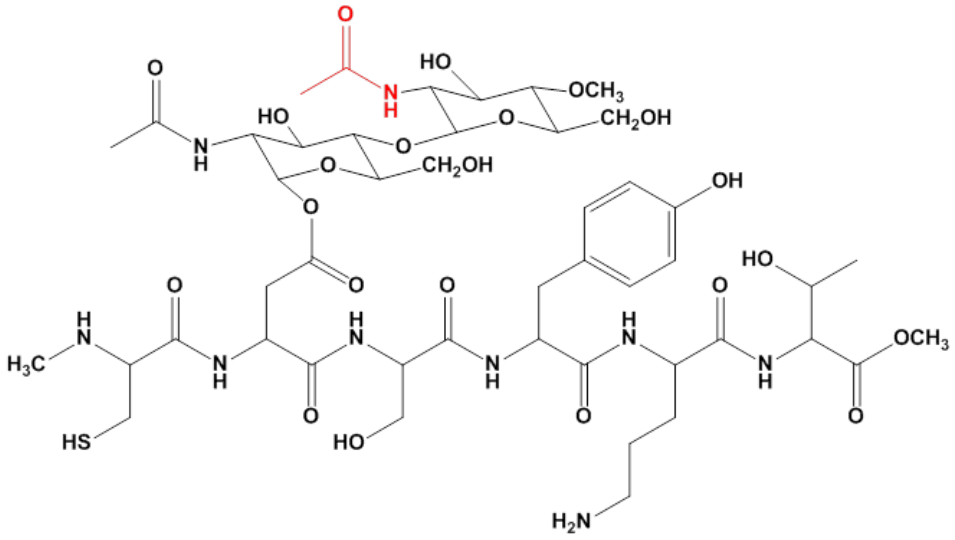


Figure S3. Representation of GlnNAc substrate. Positions highlighted in red evidence the atoms where the catalysis happens.

MD simulations

Starting from the crystallographic structure (PDB ID: 4L1G)[1] the Amber16 suite of programs was used for periodic boundary simulation and analysis,[3] with the ff99SB force-field for protein atoms,[6] TIP3P for water,[7] and ZAFF[8] for the Zn^{2+} coordinated by His131, His135, Asp81 and water molecule. The protonation states for all amino acids at physiological pH were calculated according to H++ server[9].

The protein was solvated in an orthorhombic box (81 Å x 73 Å x 88 Å) using TIP3P water molecules as the solvent model, meanwhile the counter ions (Cl-) were added within 2 Å of the protein in order to neutralize the net charge.

The SHAKE[10] algorithm was used during all MD simulations, to constrain bonds involving hydrogen atoms. A progressive heating was performed from 0 to 310 K, in 50 ps, followed by another 50 ps at 310 K, in the NVT ensemble. 80 ns NPT production simulations were performed at 310 K and 1 atm. The Particle mesh Ewald summation method[11] was employed to find the electrostatic potential and the long-range electrostatic interactions with 12 Å cut-off distance.

In order to select different conformations of the alone enzyme, we performed RMSD-based clustering of the whole trajectory of 80 ns simulation time according to the relaxed complex scheme (RCS) docking protocol. [12,13]After removing overall rotation and translations by RMS-fitting the Cα atoms' positions of the trajectory, we applied the average linkage clustering algorithm, implemented in *cpptraj*, identifying 10 significant conformations. This procedure reduced the computational time required by the RCS docking protocol. [12,13]

Molecular docking of the substrate to the protein were carried out by using AutoDock [14,15] and 10 output poses were generated. Box centroid was determined by the metal ion containing active center and a box of 20 Å size for X, Y and Z was used for grid point generation. At the end, the best docking pose was obtained according to the following geometric criteria: a distance cut-off equal to 3 Å for Zn^{2+}-$O_{sub,,}$ O_{sub}-$OH_{2\text{-}HyP}$ and Nsub-$N\varepsilon_{His225}$. (Figure S4)

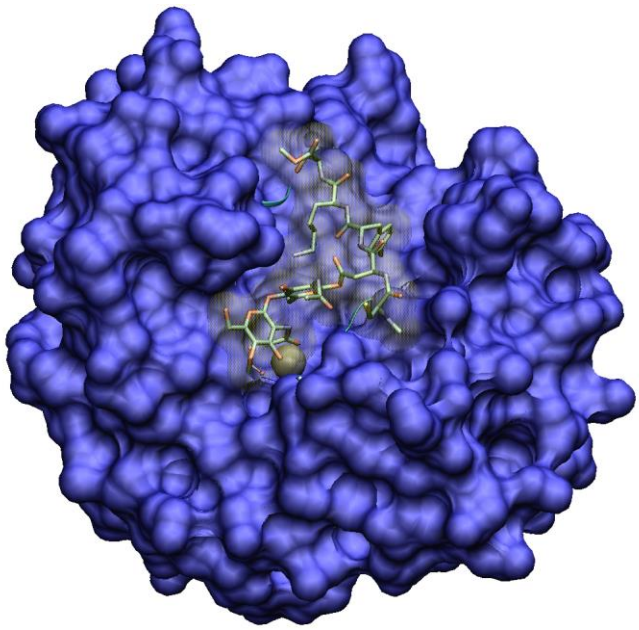


Figure S4. Best pose obtained from docking simulation.

The second step of MDs of 80 ns simulation time was applied to the substrate-bound protein with the same protocol above mentioned followed by the trajectory's clustering phase. The best enzyme-substrate complex was selected for the next QM calculations applying the same geometrical filters above mentioned. RMSD analysis was performed to monitor the simulation (Figure S5).

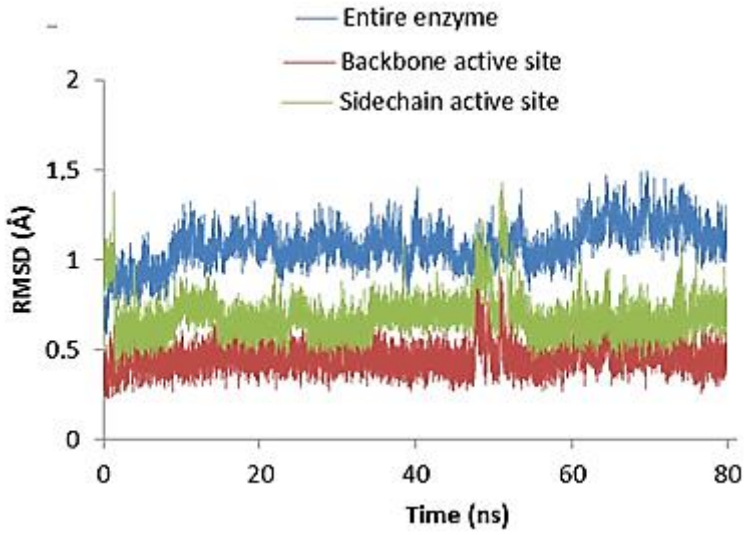


Figure S5. RMSD trend of 80 ns of PDA-2-Hyp-substrate.

Table S1. Benchmark of different hybryd DFT functionals on optimized structures of PDA-2Hyp system. All values are in kcal/mol.

	B3LYP-D3	M06	M062X	wB97XD
ES	0.0	0.0	0.0	0.0
TS1	21.7	22.8	19.5	19.4
INT	18.5	15.1	14.2	14.6
TS2	23.0	22.5	18.5	19.8
EP	-5.7	-7.9	-7.2	-8.7

References

1. V. E. Fadouloglou, S. Balomenou, M. Aivaliotis, D. Kotsifaki, S. Arnaouteli, A. Tomatsidou, G. Efstathiou, N. Kountourakis, S. Miliara, M. Griniezaki, A. Tsalafouta, S. A. Pergantis, I. G. Boneca, N. M. Glykos, V. Bouriotis, M. Kokkinidis, *JACS* **2017***, 139*, 5330-5337.
2. F. Caufrier, A. Martinou, C. Dupont, V. Bouriotis, *Carbohydr. Res.* **2003**, *338*, 687–693.
3. D. A. Case, D. S. Cerutti, T. E. III Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, (2017), AMBER **2017**, University of California, San Francisco.
4. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. Case, *J. Comp. Chem.* **2004**, *25*, 1157-1173.
5. C. I Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269-10280.
6. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712–725.
7. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
8. M. B. Peters, Y. Yang, B. Wang, L. Füsti-Molnár, M. N. Weaver, K. M. Jr Merz, *J. Chem. Theory Comput.* **2010**, *6*, 2935-2947.
9. R. Anandakrishnan, B. Aguilar, A. V. Onufriev, *Nucleic acids Res.* **2012**, *40*, W537-W541.
10. J-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327-341.
11. P. P. Ewald, *Ann Phys-Berlin* **1921**, *64*, 253-287.
12. R. E. Amaro, R. Baron, J. A. McCammon, *J. Comput. Aided Mol. Des.* **2008**, *22*, 693-705.
13. I. Romeo, N. Marascio, G. Pavia, C. Talarico, G. Costa, S. Alcaro, A. Artese, C. Torti, M. C. Liberto, A. Focà, *Chem. Sel.* **2018**, *3*, 6009-6017.
14. O. Trott, A. J. Olson, *J. Comp. Chem.* **2010**, *31*, 455-461.
15. D. Santos-Martins, F. Forli, M. J. Ramos, A. J. Olson, *J. Chem. Inf. Model.*, **2014**, *54*, 2371–2379.

ES

```
HETATM   1 C    0   -0.278   3.029   1.918          C
HETATM   2 H    0    0.090   4.053   2.045          H
HETATM   3 H    0   -0.124   2.472   2.841          H
HETATM   4 C    0    0.426   2.398   0.744          C
HETATM   5 O    0    0.199   2.719  -0.417          O
HETATM   6 O    0    1.314   1.472   1.078          O
HETATM   7 C    0    3.385  -2.008   3.510          C
HETATM   8 H    0    4.119  -2.802   3.361          H
HETATM   9 H    0    2.861  -2.143   4.459          H
HETATM  10 C    0    2.398  -1.952   2.358          C
HETATM  11 O    0    1.280  -1.365   2.586          O
HETATM  12 O    0    2.732  -2.423   1.236          O
HETATM  13 C    0   -4.149   1.756   3.005          C
HETATM  14 H    0   -4.093   2.847   2.924          H
HETATM  15 H    0   -3.980   1.487   4.049          H
HETATM  16 C    0   -3.136   1.089   2.137          C
HETATM  17 N    0   -3.131   1.220   0.752          N
HETATM  18 H    0   -3.921   1.506   0.167          H
HETATM  19 C    0   -2.099   0.493   0.264          C
HETATM  20 H    0   -1.875   0.377  -0.781          H
HETATM  21 N    0   -1.414  -0.064   1.250          N
HETATM  22 C    0   -2.052   0.300   2.426          C
HETATM  23 H    0   -1.660   0.017   3.391          H
HETATM  24 C    0   -2.566  -4.952   4.098          C
HETATM  25 H    0   -2.607  -4.210   4.899          H
HETATM  26 H    0   -2.035  -5.831   4.483          H
HETATM  27 C    0   -1.880  -4.373   2.907          C
HETATM  28 N    0   -1.679  -5.077   1.727          N
HETATM  29 H    0   -1.954  -6.034   1.556          H
HETATM  30 C    0   -1.016  -4.283   0.849          C
HETATM  31 H    0   -0.739  -4.561  -0.157          H
HETATM  32 N    0   -0.788  -3.097   1.392          N
HETATM  33 C    0   -1.314  -3.147   2.675          C
HETATM  34 H    0   -1.215  -2.313   3.351          H
HETATM  35 C    0   -3.412   4.464  -1.759          C
HETATM  36 H    0   -2.533   4.891  -1.263          H
HETATM  37 H    0   -3.820   3.645  -1.161          H
HETATM  38 N    0   -3.090   3.973  -3.099          N
HETATM  39 H    0   -3.473   4.449  -3.901          H
HETATM  40 C    0   -2.285   2.920  -3.311          C
HETATM  41 N    0   -1.547   2.417  -2.344          N
HETATM  42 H    0   -1.230   2.940  -1.527          H
HETATM  43 H    0   -1.069   1.500  -2.468          H
HETATM  44 N    0   -2.231   2.359  -4.545          N
HETATM  45 H    0   -3.017   2.418  -5.173          H
HETATM  46 H    0   -1.561   1.617  -4.695          H
HETATM  47 N    0   -4.692  -0.912  -1.103          N
HETATM  48 C    0   -3.989  -2.154  -1.432          C
HETATM  49 O    0   -2.630  -1.883  -1.799          O
HETATM  50 H    0   -2.008  -2.488  -1.359          H
HETATM  51 C    0   -4.717  -2.936  -2.551          C
HETATM  52 O    0   -5.937  -3.051  -2.516          O
HETATM  53 C    0   -4.106  -2.968  -0.120          C
HETATM  54 H    0   -3.273  -2.681   0.530          H
HETATM  55 H    0   -4.053  -4.043  -0.308          H
HETATM  56 C    0   -5.429  -2.493   0.479          C
HETATM  57 H    0   -5.517  -2.726   1.543          H
HETATM  58 H    0   -6.268  -2.932  -0.064          H
HETATM  59 C    0   -5.354  -0.987   0.222          C
HETATM  60 H    0   -4.758  -0.508   0.999          H
HETATM  61 H    0   -6.322  -0.486   0.166          H
HETATM  62 C    0    9.723   3.642   0.691          C
HETATM  63 H    0   10.558   3.160   0.183          H
HETATM  64 H    0    9.701   3.340   1.743          H
HETATM  65 C    0    8.418   3.260   0.018          C
HETATM  66 O    0    7.365   3.835   0.611          O
HETATM  67 O    0    8.335   2.495  -0.927          O
HETATM  68 C    0    4.655   2.202   2.221          C
HETATM  69 H    0    3.935   3.020   2.341          H
HETATM  70 H    0    5.644   2.592   2.481          H
HETATM  71 C    0    4.657   1.697   0.818          C
HETATM  72 N    0    5.195   2.464  -0.200          N
HETATM  73 H    0    5.971   3.050   0.032          H
HETATM  74 C    0    5.083   1.751  -1.307          C
HETATM  75 H    0    5.457   2.037  -2.279          H
HETATM  76 N    0    4.458   0.574  -1.061          N
HETATM  77 H    0    4.364  -0.243  -1.657          H
HETATM  78 C    0    4.199   0.513   0.296          C
HETATM  79 H    0    3.734  -0.354   0.737          H
HETATM  80 Zn   0    0.225  -1.292   0.881         Zn
HETATM  81 C    0    4.223  -4.947  -0.050          C
HETATM  82 O    0    5.511  -5.484  -0.342          O
HETATM  83 C    0    6.413  -4.564  -1.003          C
HETATM  84 C    0    6.920  -3.501  -0.007          C
HETATM  85 H    0    6.116  -2.798   0.288          H
HETATM  86 H    0    7.702  -2.892  -0.483          H
HETATM  87 O    0    7.487  -4.129   1.123          O
HETATM  88 H    0    6.904  -4.883   1.330          H
HETATM  89 H    0    7.280  -5.194  -1.253          H
HETATM  90 H    0    4.267  -4.115   0.668          H
HETATM  91 C    0    3.506  -4.504  -1.351          C
HETATM  92 N    0    2.157  -3.995  -1.079          N
HETATM  93 H    0    2.044  -3.454  -0.214          H
HETATM  94 C    0    1.132  -3.995  -1.964          C
HETATM  95 O    0    0.034  -3.451  -1.712          O
HETATM  96 C    0    1.319  -4.752  -3.265          C
HETATM  97 H    0    1.417  -5.829  -3.073          H
HETATM  98 H    0    2.214  -4.438  -3.811          H
HETATM  99 H    0    0.440  -4.580  -3.886          H
HETATM 100 H    0    3.454  -5.393  -1.989          H
HETATM 101 C    0    4.364  -3.472  -2.109          C
HETATM 102 O    0    4.358  -2.172  -1.529          O
HETATM 103 H    0    4.324  -2.188  -0.534          H
HETATM 104 H    0    3.957  -3.324  -3.112          H
HETATM 105 C    0    5.802  -4.021  -2.301          C
HETATM 106 H    0    6.433  -3.199  -2.685          H
HETATM 107 O    0    5.655  -5.052  -3.276          O
HETATM 108 C    0    6.847  -5.392  -3.976          C
HETATM 109 H    0    6.574  -6.156  -4.712          H
HETATM 110 H    0    7.622  -5.809  -3.317          H
HETATM 111 H    0    7.269  -4.518  -4.499          H
HETATM 112 O    0   -0.524  -0.068  -2.715          O
HETATM 113 H    0    0.157  -0.288  -2.003          H
HETATM 114 H    0   -1.216  -0.737  -2.620          H
HETATM 115 O    0    1.100  -0.582  -0.670          O
HETATM 116 H    0    1.632   0.974   0.281          H
HETATM 117 H    0    1.752  -1.255  -0.911          H
HETATM 118 H    0    9.848   4.729   0.670          H
HETATM 119 H    0   -3.594  -5.256   3.867          H
HETATM 120 H    0   -5.169   1.456   2.738          H
HETATM 121 H    0    3.918  -1.048   3.557          H
HETATM 122 H    0   -1.345   3.099   1.696          H
HETATM 123 H    0   -4.169   5.242  -1.854          H
HETATM 124 H    0    4.386   1.410   2.924          H
HETATM 125 C    0   -4.851   0.205  -1.884          C
HETATM 126 C    0   -4.775   0.090  -3.399          C
HETATM 127 H    0   -3.729   0.130  -3.722          H
HETATM 128 H    0   -5.212  -0.842  -3.760          H
HETATM 129 H    0   -5.326   0.939  -3.845          H
HETATM 130 O    0   -5.088   1.302  -1.354          O
HETATM 131 N    0   -3.903  -3.527  -3.456          N
HETATM 132 H    0   -2.918  -3.307  -3.404          H
HETATM 133 C    0   -4.408  -4.372  -4.527          C
HETATM 134 H    0   -3.688  -4.440  -5.350          H
HETATM 135 H    0   -5.330  -3.926  -4.901          H
HETATM 136 O    0    3.539  -6.561   0.450          O
HETATM 137 C    0    3.258  -6.652   1.849          C
HETATM 138 C    0    4.632  -7.224   2.500          C
HETATM 139 H    0    3.048  -5.676   2.235          H
HETATM 140 C    0    1.751  -6.833   2.637          C
HETATM 141 C    0    4.329  -7.322   4.006          C
HETATM 142 H    0    4.841  -8.200   2.114          H
HETATM 143 O    0    5.766  -6.378   2.296          O
HETATM 144 H    0    1.621  -7.631   1.935          H
HETATM 145 O    0    1.816  -7.410   4.063          O
HETATM 146 C    0    0.810  -5.615   2.618          C
HETATM 147 H    0    4.119  -6.346   4.392          H
HETATM 148 C    0    3.107  -8.234   4.226          C
HETATM 149 N    0    5.493  -7.886   4.705          N
HETATM 150 H    0    5.823  -6.135   1.369          H
HETATM 151 H    0    0.753  -5.223   1.624          H
HETATM 152 H    0   -0.166  -5.913   2.939          H
HETATM 153 O    0    1.315  -4.608   3.499          O
HETATM 154 H    0    3.144  -8.648   5.212          H
HETATM 155 H    0    3.118  -9.025   3.505          H
HETATM 156 H    0    5.897  -7.188   5.296          H
HETATM 157 C    0    6.594  -8.915   4.142          C
HETATM 158 H    0    0.696  -3.876   3.538          H
HETATM 159 C    0    7.925  -9.243   4.845          C
HETATM 160 O    0    6.248  -9.597   3.143          O
HETATM 161 H    0    8.741  -8.983   4.203          H
HETATM 162 H    0    7.995  -8.686   5.755          H
HETATM 163 H    0    7.964 -10.290   5.064          H
HETATM 164 C    0   -4.702  -5.791  -4.006          C
HETATM 165 H    0   -5.686  -5.821  -3.588          H
HETATM 166 H    0   -4.636  -6.489  -4.814          H
HETATM 167 C    0   -3.673  -6.163  -2.922          C
HETATM 168 C    0   -4.104  -6.662  -1.692          C
HETATM 169 C    0   -2.310  -6.003  -3.170          C
HETATM 170 C    0   -3.173  -7.000  -0.711          C
HETATM 171 H    0   -5.179  -6.788  -1.497          H
HETATM 172 C    0   -1.378  -6.340  -2.188          C
HETATM 173 H    0   -1.969  -5.610  -4.139          H
HETATM 174 C    0   -1.809  -6.838  -0.959          C
HETATM 175 H    0   -3.513  -7.393   0.258          H
HETATM 176 H    0   -0.303  -6.214  -2.384          H
HETATM 177 O    0   -0.854  -7.184   0.048          O
HETATM 178 H    0   -0.581  -8.097  -0.070          H
```

TS1

```
HETATM   1 C    0   -0.278   3.029   1.918           C
HETATM   2 H    0    0.144   3.973   2.280           H
HETATM   3 H    0   -0.372   2.363   2.774           H
HETATM   4 C    0    0.663   2.447   0.906           C
HETATM   5 O    0    0.484   2.425  -0.306           O
HETATM   6 O    0    1.744   1.973   1.506           O
HETATM   7 C    0    2.528  -2.004   4.345           C
HETATM   8 H    0    2.795  -2.963   4.792           H
HETATM   9 H    0    1.912  -1.415   5.028           H
HETATM  10 C    0    1.820  -2.213   3.019           C
HETATM  11 O    0    1.122  -1.205   2.597           O
HETATM  12 O    0    1.969  -3.278   2.389           O
HETATM  13 C    0   -3.897   1.966   3.385           C
HETATM  14 H    0   -3.719   3.047   3.431           H
HETATM  15 H    0   -3.733   1.557   4.384           H
HETATM  16 C    0   -2.990   1.302   2.405           C
HETATM  17 N    0   -3.032   1.568   1.040           N
HETATM  18 H    0   -3.825   1.945   0.522           H
HETATM  19 C    0   -2.095   0.801   0.429           C
HETATM  20 H    0   -1.915   0.793  -0.634           H
HETATM  21 N    0   -1.429   0.085   1.318           N
HETATM  22 C    0   -1.975   0.392   2.558           C
HETATM  23 H    0   -1.563  -0.033   3.460           H
HETATM  24 C    0   -3.120  -4.771   4.776           C
HETATM  25 H    0   -3.727  -3.971   5.207           H
HETATM  26 H    0   -2.502  -5.189   5.578           H
HETATM  27 C    0   -2.286  -4.230   3.663           C
HETATM  28 N    0   -1.416  -5.013   2.917           N
HETATM  29 H    0   -1.232  -5.995   3.062           H
HETATM  30 C    0   -0.809  -4.232   1.992           C
HETATM  31 H    0   -0.069  -4.561   1.283           H
HETATM  32 N    0   -1.242  -2.985   2.091           N
HETATM  33 C    0   -2.159  -2.973   3.132           C
HETATM  34 H    0   -2.678  -2.074   3.421           H
HETATM  35 C    0   -3.412   4.464  -1.759           C
HETATM  36 H    0   -2.669   5.182  -2.114           H
HETATM  37 H    0   -3.206   4.235  -0.705           H
HETATM  38 N    0   -3.387   3.276  -2.597           N
HETATM  39 H    0   -4.207   2.670  -2.544           H
HETATM  40 C    0   -2.253   2.702  -3.023           C
HETATM  41 N    0   -1.060   3.067  -2.536           N
HETATM  42 H    0   -0.931   3.253  -1.543           H
HETATM  43 H    0   -0.253   2.602  -2.928           H
HETATM  44 N    0   -2.299   1.788  -4.003           N
HETATM  45 H    0   -3.158   1.648  -4.511           H
HETATM  46 H    0   -1.614   1.035  -3.939           H
HETATM  47 N    0   -4.467  -0.936  -1.104           N
HETATM  48 C    0   -3.854  -2.227  -1.489           C
HETATM  49 O    0   -2.503  -2.065  -1.935           O
HETATM  50 H    0   -1.835  -2.532  -1.375           H
HETATM  51 C    0   -4.717  -2.936  -2.551           C
HETATM  52 O    0   -5.940  -2.940  -2.436           O
HETATM  53 C    0   -3.950  -3.071  -0.198           C
HETATM  54 H    0   -3.036  -2.917   0.380           H
HETATM  55 H    0   -4.036  -4.136  -0.427           H
HETATM  56 C    0   -5.157  -2.489   0.536           C
HETATM  57 H    0   -5.153  -2.739   1.600           H
HETATM  58 H    0   -6.086  -2.834   0.078           H
HETATM  59 C    0   -4.959  -0.994   0.298           C
HETATM  60 H    0   -4.214  -0.616   0.995           H
HETATM  61 H    0   -5.863  -0.389   0.385           H
HETATM  62 C    0    9.723   3.642   0.691           C
HETATM  63 H    0   10.713   3.264   0.441           H
HETATM  64 H    0    9.627   3.760   1.775           H
HETATM  65 C    0    8.666   2.683   0.182           C
HETATM  66 O    0    7.427   3.107   0.477           O
HETATM  67 O    0    8.900   1.653  -0.420           O
HETATM  68 C    0    4.655   2.202   2.221           C
HETATM  69 H    0    4.264   3.161   1.866           H
HETATM  70 H    0    5.685   2.377   2.544           H
HETATM  71 C    0    4.630   1.169   1.131           C
HETATM  72 N    0    5.462   1.251   0.031           N
HETATM  73 H    0    6.206   1.916   0.101           H
HETATM  74 C    0    5.202   0.192  -0.719           C
HETATM  75 H    0    5.685  -0.061  -1.651           H
HETATM  76 N    0    4.227  -0.575  -0.167           N
HETATM  77 H    0    3.841  -1.463  -0.517           H
HETATM  78 C    0    3.852   0.033   1.020           C
HETATM  79 H    0    3.102  -0.389   1.672           H
HETATM  80 Zn   0   -0.068  -1.497   1.020          Zn
HETATM  81 C    0    4.158  -4.589   0.103           C
HETATM  82 O    0    4.785  -5.806  -0.276           O
HETATM  83 C    0    5.578  -5.742  -1.471           C
HETATM  84 C    0    6.946  -5.124  -1.152           C
HETATM  85 H    0    6.861  -4.051  -0.894           H
HETATM  86 H    0    7.593  -5.181  -2.034           H
HETATM  87 O    0    7.572  -5.835  -0.110           O
HETATM  88 H    0    6.886  -6.001   0.554           H
HETATM  89 H    0    5.771  -6.792  -1.710           H
HETATM  90 H    0    4.393  -3.753   0.781           H
HETATM  91 C    0    3.181  -4.107  -1.003           C
HETATM  92 N    0    2.437  -2.918  -0.520           N
HETATM  93 H    0    2.249  -3.047   0.484           H
HETATM  94 C    0    1.152  -2.541  -1.117           C
HETATM  95 O    0    0.039  -3.003  -0.548           O
HETATM  96 C    0    1.165  -2.530  -2.636           C
HETATM  97 H    0    1.212  -3.564  -2.988           H
HETATM  98 H    0    2.017  -1.977  -3.037           H
HETATM  99 H    0    0.243  -2.085  -3.005           H
HETATM 100 H    0    2.515  -4.950  -1.232           H
HETATM 101 C    0    4.030  -3.837  -2.267           C
HETATM 102 O    0    4.874  -2.694  -2.134           O
HETATM 103 H    0    5.592  -2.888  -1.517           H
HETATM 104 H    0    3.395  -3.617  -3.123           H
HETATM 105 C    0    4.813  -5.110  -2.648           C
HETATM 106 H    0    5.530  -4.837  -3.439           H
HETATM 107 O    0    3.826  -6.002  -3.155           O
HETATM 108 C    0    4.320  -6.997  -4.036           C
HETATM 109 H    0    3.460  -7.583  -4.366           H
HETATM 110 H    0    5.034  -7.676  -3.548           H
HETATM 111 H    0    4.810  -6.550  -4.914           H
HETATM 112 O    0   -0.801   0.043  -2.560           O
HETATM 113 H    0   -0.044  -0.145  -1.967           H
HETATM 114 H    0   -1.435  -0.683  -2.393           H
HETATM 115 O    0    1.037  -0.958  -0.662           O
HETATM 116 H    0    2.392   1.586   0.876           H
HETATM 117 H    0    1.890  -0.493  -0.615           H
HETATM 118 H    0    9.574   4.629   0.243           H
HETATM 119 H    0   -3.798  -5.559   4.431           H
HETATM 120 H    0   -4.950   1.813   3.125           H
HETATM 121 H    0    3.451  -1.442   4.162           H
HETATM 122 H    0   -1.251   3.223   1.470           H
HETATM 123 H    0   -4.401   4.918  -1.829           H
HETATM 124 H    0    4.055   1.884   3.076           H
HETATM 125 C    0   -4.800   0.130  -1.895           C
HETATM 126 C    0   -5.006  -0.027  -3.399           C
HETATM 127 H    0   -4.122  -0.427  -3.904           H
HETATM 128 H    0   -5.811  -0.752  -3.553           H
HETATM 129 H    0   -5.326   0.940  -3.845           H
HETATM 130 O    0   -5.022   1.245  -1.378           O
HETATM 131 N    0   -4.023  -3.556  -3.529           N
HETATM 132 H    0   -3.019  -3.447  -3.516           H
HETATM 133 C    0   -4.661  -4.364  -4.558           C
HETATM 134 H    0   -3.973  -4.556  -5.387           H
HETATM 135 H    0   -5.525  -3.813  -4.930           H
HETATM 136 O    0    3.552  -5.610   0.901           O
HETATM 137 C    0    3.772  -5.925   2.509           C
HETATM 138 C    0    4.908  -6.086   3.660           C
HETATM 139 C    0    2.364  -6.896   2.499           C
HETATM 140 C    0    4.163  -6.729   4.844           C
HETATM 141 H    0    5.680  -6.746   3.323           H
HETATM 142 O    0    5.489  -4.848   4.078           O
HETATM 143 H    0    2.895  -7.438   1.745           H
HETATM 144 O    0    2.171  -7.792   3.737           O
HETATM 145 C    0    1.038  -6.187   2.167           C
HETATM 146 H    0    3.391  -6.069   5.181           H
HETATM 147 C    0    3.538  -8.062   4.393           C
HETATM 148 N    0    5.109  -6.978   5.941           N
HETATM 149 H    0    5.756  -4.342   3.307           H
HETATM 150 H    0    1.169  -5.575   1.299           H
HETATM 151 H    0    0.281  -6.919   1.976           H
HETATM 152 O    0    0.644  -5.369   3.271           O
HETATM 153 H    0    3.405  -8.699   5.243           H
HETATM 154 H    0    4.185  -8.539   3.687           H
HETATM 155 H    0    4.887  -6.384   6.714           H
HETATM 156 C    0    6.685  -7.288   5.851           C
HETATM 157 H    0   -0.214  -4.979   3.091           H
HETATM 158 C    0    7.652  -7.268   7.050           C
HETATM 159 O    0    7.095  -7.726   4.745           O
HETATM 160 H    0    8.415  -6.537   6.880           H
HETATM 161 H    0    7.111  -7.019   7.939           H
HETATM 162 H    0    8.100  -8.232   7.163           H
HETATM 163 H    0    3.405  -4.976   2.839           H
HETATM 164 C    0   -5.129  -5.713  -3.980           C
HETATM 165 H    0   -6.102  -5.599  -3.551           H
HETATM 166 H    0   -5.166  -6.442  -4.762           H
HETATM 167 C    0   -4.141  -6.175  -2.893           C
HETATM 168 C    0   -4.618  -6.603  -1.654           C
HETATM 169 C    0   -2.769  -6.168  -3.148           C
HETATM 170 C    0   -3.723  -7.024  -0.669           C
HETATM 171 H    0   -5.699  -6.609  -1.453           H
HETATM 172 C    0   -1.874  -6.588  -2.163           C
HETATM 173 H    0   -2.393  -5.832  -4.125           H
HETATM 174 C    0   -2.351  -7.016  -0.924           C
HETATM 175 H    0   -4.099  -7.361   0.307           H
HETATM 176 H    0   -0.793  -6.582  -2.364           H
HETATM 177 O    0   -1.435  -7.446   0.086           O
HETATM 178 H    0   -1.263  -8.385  -0.017           H
```

INT

```
HETATM   1 C     0    -0.278   3.029   1.918           C
HETATM   2 H     0     0.267   3.915   2.260           H
HETATM   3 H     0    -0.510   2.422   2.793           H
HETATM   4 C     0     0.600   2.264   0.962           C
HETATM   5 O     0     0.452   2.253  -0.251           O
HETATM   6 O     0     1.581   1.625   1.596           O
HETATM   7 C     0     2.697  -1.324   4.291           C
HETATM   8 H     0     3.684  -1.762   4.443           H
HETATM   9 H     0     2.058  -1.506   5.158           H
HETATM  10 C     0     2.055  -1.869   3.030           C
HETATM  11 O     0     0.801  -1.582   2.888           O
HETATM  12 O     0     2.731  -2.499   2.195           O
HETATM  13 C     0    -3.770   1.897   2.861           C
HETATM  14 H     0    -3.441   2.942   2.822           H
HETATM  15 H     0    -3.771   1.585   3.908           H
HETATM  16 C     0    -2.875   1.011   2.063           C
HETATM  17 N     0    -2.762   1.115   0.681           N
HETATM  18 H     0    -3.442   1.553   0.057           H
HETATM  19 C     0    -1.886   0.175   0.257           C
HETATM  20 H     0    -1.628   0.009  -0.771           H
HETATM  21 N     0    -1.406  -0.506   1.281           N
HETATM  22 C     0    -2.008   0.009   2.420           C
HETATM  23 H     0    -1.739  -0.355   3.399           H
HETATM  24 C     0    -3.683  -5.895   3.744           C
HETATM  25 H     0    -4.074  -5.211   4.500           H
HETATM  26 H     0    -3.226  -6.743   4.266           H
HETATM  27 C     0    -2.696  -5.185   2.880           C
HETATM  28 N     0    -2.051  -5.791   1.811           N
HETATM  29 H     0    -2.161  -6.754   1.528           H
HETATM  30 C     0    -1.213  -4.891   1.238           C
HETATM  31 H     0    -0.587  -5.066   0.376           H
HETATM  32 N     0    -1.290  -3.737   1.878           N
HETATM  33 C     0    -2.203  -3.905   2.901           C
HETATM  34 H     0    -2.440  -3.102   3.583           H
HETATM  35 C     0    -3.412   4.464  -1.759           C
HETATM  36 H     0    -2.602   4.785  -1.096           H
HETATM  37 H     0    -4.000   3.670  -1.291           H
HETATM  38 N     0    -2.892   3.998  -3.047           N
HETATM  39 H     0    -3.215   4.448  -3.889           H
HETATM  40 C     0    -2.142   2.894  -3.173           C
HETATM  41 N     0    -1.544   2.340  -2.126           N
HETATM  42 H     0    -1.200   2.861  -1.326           H
HETATM  43 H     0    -1.058   1.448  -2.307           H
HETATM  44 N     0    -1.987   2.314  -4.381           N
HETATM  45 H     0    -2.665   2.456  -5.113           H
HETATM  46 H     0    -1.379   1.502  -4.437           H
HETATM  47 N     0    -4.479  -0.890  -1.151           N
HETATM  48 C     0    -3.851  -2.172  -1.516           C
HETATM  49 O     0    -2.512  -2.012  -2.001           O
HETATM  50 H     0    -1.843  -2.354  -1.368           H
HETATM  51 C     0    -4.717  -2.936  -2.551           C
HETATM  52 O     0    -5.939  -2.927  -2.447           O
HETATM  53 C     0    -3.896  -2.957  -0.187           C
HETATM  54 H     0    -3.011  -2.684   0.396           H
HETATM  55 H     0    -3.874  -4.036  -0.355           H
HETATM  56 C     0    -5.160  -2.446   0.501           C
HETATM  57 H     0    -5.155  -2.645   1.576           H
HETATM  58 H     0    -6.045  -2.897   0.050           H
HETATM  59 C     0    -5.106  -0.947   0.194           C
HETATM  60 H     0    -4.504  -0.431   0.941           H
HETATM  61 H     0    -6.085  -0.466   0.149           H
HETATM  62 C     0     9.723   3.642   0.691           C
HETATM  63 H     0    10.638   3.191   0.310           H
HETATM  64 H     0     9.680   3.541   1.780           H
HETATM  65 C     0     8.519   2.964   0.067           C
HETATM  66 O     0     7.368   3.512   0.485           O
HETATM  67 O     0     8.581   2.035  -0.717           O
HETATM  68 C     0     4.655   2.202   2.221           C
HETATM  69 H     0     3.973   3.051   2.113           H
HETATM  70 H     0     5.653   2.603   2.419           H
HETATM  71 C     0     4.689   1.366   0.975           C
HETATM  72 N     0     5.330   1.831  -0.158           N
HETATM  73 H     0     6.066   2.492  -0.012           H
HETATM  74 C     0     5.258   0.863  -1.053           C
HETATM  75 H     0     5.707   0.885  -2.036           H
HETATM  76 N     0     4.564  -0.201  -0.577           N
HETATM  77 H     0     4.468  -1.113  -1.034           H
HETATM  78 C     0     4.198   0.098   0.729           C
HETATM  79 H     0     3.739  -0.648   1.362           H
HETATM  80 Zn    0    -0.104  -2.120   1.258          Zn
HETATM  81 C     0     4.158  -4.589   0.103           C
HETATM  82 O     0     4.745  -5.826  -0.286           O
HETATM  83 C     0     5.625  -5.767  -1.420           C
HETATM  84 C     0     7.001  -5.247  -0.981           C
HETATM  85 H     0     6.963  -4.178  -0.695           H
HETATM  86 H     0     7.708  -5.319  -1.814           H
HETATM  87 O     0     7.501  -6.024   0.081           O
HETATM  88 H     0     6.754  -6.168   0.683           H
HETATM  89 H     0     5.774  -6.819  -1.686           H
HETATM  90 H     0     4.909  -3.850   0.422           H
HETATM  91 C     0     3.294  -4.022  -1.054           C
HETATM  92 N     0     2.555  -2.806  -0.680           N
HETATM  93 H     0     2.502  -2.707   0.338           H
HETATM  94 C     0     1.243  -2.494  -1.253           C
HETATM  95 O     0     0.151  -2.866  -0.501           O
HETATM  96 C     0     1.067  -3.002  -2.687           C
HETATM  97 H     0     1.147  -4.089  -2.739           H
HETATM  98 H     0     1.803  -2.560  -3.362           H
HETATM  99 H     0     0.072  -2.714  -3.032           H
HETATM 100 H     0     2.632  -4.845  -1.350           H
HETATM 101 C     0     4.246  -3.756  -2.242           C
HETATM 102 O     0     5.146  -2.671  -1.991           O
HETATM 103 H     0     5.802  -2.937  -1.332           H
HETATM 104 H     0     3.681  -3.449  -3.121           H
HETATM 105 C     0     4.989  -5.050  -2.626           C
HETATM 106 H     0     5.777  -4.791  -3.352           H
HETATM 107 O     0     3.996  -5.866  -3.239           O
HETATM 108 C     0     4.502  -6.848  -4.127           C
HETATM 109 H     0     3.640  -7.372  -4.545           H
HETATM 110 H     0     5.143  -7.585  -3.623           H
HETATM 111 H     0     5.076  -6.392  -4.948           H
HETATM 112 O     0    -0.672  -0.083  -3.042           O
HETATM 113 H     0     0.100  -0.390  -2.498           H
HETATM 114 H     0    -1.373  -0.737  -2.868           H
HETATM 115 O     0     1.195  -1.011  -1.298           O
HETATM 116 H     0     2.197   1.200   0.960           H
HETATM 117 H     0     2.117  -0.711  -1.328           H
HETATM 118 H     0     9.706   4.712   0.464           H
HETATM 119 H     0    -4.531  -6.275   3.163           H
HETATM 120 H     0    -4.802   1.858   2.495           H
HETATM 121 H     0     2.796  -0.238   4.181           H
HETATM 122 H     0    -1.194   3.340   1.417           H
HETATM 123 H     0    -4.055   5.324  -1.944           H
HETATM 124 H     0     4.327   1.613   3.082           H
HETATM 125 C     0    -4.657   0.232  -1.921           C
HETATM 126 C     0    -4.673   0.139  -3.442           C
HETATM 127 H     0    -3.662   0.282  -3.836           H
HETATM 128 H     0    -5.048  -0.824  -3.789           H
HETATM 129 H     0    -5.326   0.939  -3.845           H
HETATM 130 O     0    -4.844   1.328  -1.367           O
HETATM 131 N     0    -4.016  -3.626  -3.479           N
HETATM 132 H     0    -3.016  -3.491  -3.487           H
HETATM 133 C     0    -4.654  -4.456  -4.489           C
HETATM 134 H     0    -3.986  -4.627  -5.340           H
HETATM 135 H     0    -5.546  -3.936  -4.838           H
HETATM 136 O     0     3.349  -4.863   1.250           O
HETATM 137 C     0     3.284  -6.288   1.960           C
HETATM 138 C     0     4.738  -6.967   2.215           C
HETATM 139 H     0     2.839  -6.014   2.894           H
HETATM 140 C     0     1.844  -7.209   2.018           C
HETATM 141 C     0     4.429  -8.237   3.028           C
HETATM 142 H     0     5.183  -7.240   1.281           H
HETATM 143 O     0     5.644  -6.139   2.949           O
HETATM 144 H     0     1.925  -7.172   0.952           H
HETATM 145 O     0     2.009  -8.664   2.496           O
HETATM 146 C     0     0.641  -6.566   2.734           C
HETATM 147 H     0     3.985  -7.963   3.962           H
HETATM 148 C     0     3.454  -9.129   2.237           C
HETATM 149 N     0     5.676  -8.975   3.277           N
HETATM 150 H     0     5.664  -5.262   2.560           H
HETATM 151 H     0     0.518  -5.560   2.392           H
HETATM 152 H     0    -0.244  -7.127   2.517           H
HETATM 153 O     0     0.871  -6.565   4.145           O
HETATM 154 H     0     3.563 -10.145   2.553           H
HETATM 155 H     0     3.671  -9.055   1.192           H
HETATM 156 H     0     5.888  -8.951   4.254           H
HETATM 157 C     0     7.004  -9.028   2.370           C
HETATM 158 H     0     0.095  -6.225   4.596           H
HETATM 159 C     0     8.357  -9.602   2.831           C
HETATM 160 O     0     6.854  -8.717   1.160           O
HETATM 161 H     0     9.103  -8.837   2.787           H
HETATM 162 H     0     8.269  -9.957   3.837           H
HETATM 163 H     0     8.638 -10.412   2.191           H
HETATM 164 C     0    -5.060  -5.818  -3.896           C
HETATM 165 H     0    -6.020  -5.733  -3.432           H
HETATM 166 H     0    -5.104  -6.548  -4.676           H
HETATM 167 C     0    -4.020  -6.251  -2.846           C
HETATM 168 C     0    -4.437  -6.669  -1.582           C
HETATM 169 C     0    -2.661  -6.226  -3.159           C
HETATM 170 C     0    -3.495  -7.062  -0.631           C
HETATM 171 H     0    -5.509  -6.689  -1.336           H
HETATM 172 C     0    -1.718  -6.619  -2.207           C
HETATM 173 H     0    -2.331  -5.897  -4.155           H
HETATM 174 C     0    -2.136  -7.037  -0.944           C
HETATM 175 H     0    -3.824  -7.392   0.365           H
HETATM 176 H     0    -0.647  -6.599  -2.454           H
HETATM 177 O     0    -1.171  -7.439   0.032           O
HETATM 178 H     0    -0.988  -8.377  -0.066           H
```

TS2

HETATM	1	C	0	-0.278	3.029	1.918	C
HETATM	2	H	0	0.151	3.863	2.477	H
HETATM	3	H	0	-0.420	2.198	2.615	H
HETATM	4	C	0	0.662	2.575	0.839	C
HETATM	5	O	0	0.319	2.216	-0.282	O
HETATM	6	O	0	1.907	2.576	1.276	O
HETATM	7	C	0	3.139	-2.021	3.896	C
HETATM	8	H	0	3.634	-2.929	4.242	H
HETATM	9	H	0	2.706	-1.469	4.732	H
HETATM	10	C	0	2.087	-2.347	2.863	C
HETATM	11	O	0	1.193	-1.441	2.648	O
HETATM	12	O	0	2.130	-3.423	2.219	O
HETATM	13	C	0	-3.987	1.937	2.797	C
HETATM	14	H	0	-3.673	2.987	2.774	H
HETATM	15	H	0	-4.054	1.631	3.843	H
HETATM	16	C	0	-3.023	1.064	2.067	C
HETATM	17	N	0	-2.829	1.156	0.694	N
HETATM	18	H	0	-3.487	1.574	0.031	H
HETATM	19	C	0	-1.909	0.235	0.332	C
HETATM	20	H	0	-1.580	0.065	-0.676	H
HETATM	21	N	0	-1.481	-0.431	1.391	N
HETATM	22	C	0	-2.166	0.080	2.487	C
HETATM	23	H	0	-1.973	-0.282	3.485	H
HETATM	24	C	0	-3.842	-5.244	4.329	C
HETATM	25	H	0	-4.639	-4.498	4.369	H
HETATM	26	H	0	-3.569	-5.498	5.359	H
HETATM	27	C	0	-2.679	-4.701	3.569	C
HETATM	28	N	0	-1.511	-5.418	3.352	N
HETATM	29	H	0	-1.322	-6.352	3.688	H
HETATM	30	C	0	-0.652	-4.655	2.637	C
HETATM	31	H	0	0.355	-4.931	2.361	H
HETATM	32	N	0	-1.206	-3.482	2.374	N
HETATM	33	C	0	-2.464	-3.496	2.952	C
HETATM	34	H	0	-3.119	-2.641	2.893	H
HETATM	35	C	0	-3.412	4.464	-1.759	C
HETATM	36	H	0	-2.617	4.756	-1.064	H
HETATM	37	H	0	-4.004	3.644	-1.343	H
HETATM	38	N	0	-2.868	4.063	-3.057	N
HETATM	39	H	0	-3.101	4.604	-3.875	H
HETATM	40	C	0	-2.142	2.948	-3.213	C
HETATM	41	N	0	-1.643	2.305	-2.166	N
HETATM	42	H	0	-1.320	2.771	-1.323	H
HETATM	43	H	0	-1.149	1.421	-2.375	H
HETATM	44	N	0	-1.927	2.443	-4.445	N
HETATM	45	H	0	-2.526	2.678	-5.220	H
HETATM	46	H	0	-1.317	1.637	-4.529	H
HETATM	47	N	0	-4.493	-0.885	-1.154	N
HETATM	48	C	0	-3.853	-2.160	-1.521	C
HETATM	49	O	0	-2.514	-1.982	-2.000	O
HETATM	50	H	0	-1.851	-2.324	-1.360	H
HETATM	51	C	0	-4.717	-2.936	-2.551	C
HETATM	52	O	0	-5.937	-2.935	-2.439	O
HETATM	53	C	0	-3.887	-2.950	-0.194	C
HETATM	54	H	0	-3.005	-2.667	0.390	H
HETATM	55	H	0	-3.850	-4.027	-0.365	H
HETATM	56	C	0	-5.164	-2.458	0.487	C
HETATM	57	H	0	-5.171	-2.659	1.561	H
HETATM	58	H	0	-6.036	-2.928	0.029	H
HETATM	59	C	0	-5.140	-0.958	0.180	C
HETATM	60	H	0	-4.563	-0.425	0.934	H
HETATM	61	H	0	-6.130	-0.502	0.120	H
HETATM	62	C	0	9.723	3.642	0.691	C
HETATM	63	H	0	10.740	3.336	0.452	H
HETATM	64	H	0	9.620	3.802	1.769	H
HETATM	65	C	0	8.747	2.567	0.229	C
HETATM	66	O	0	7.476	2.885	0.520	O
HETATM	67	O	0	9.079	1.546	-0.340	O
HETATM	68	C	0	4.655	2.202	2.221	C
HETATM	69	H	0	4.505	3.215	1.834	H
HETATM	70	H	0	5.667	2.169	2.635	H
HETATM	71	C	0	4.503	1.171	1.132	C
HETATM	72	N	0	5.388	1.084	0.076	N
HETATM	73	H	0	6.191	1.675	0.152	H
HETATM	74	C	0	4.935	0.085	-0.681	C
HETATM	75	H	0	5.407	-0.232	-1.602	H
HETATM	76	N	0	3.800	-0.504	-0.208	N
HETATM	77	H	0	3.191	-1.864	-0.448	H
HETATM	78	C	0	3.528	0.198	0.962	C
HETATM	79	H	0	2.688	-0.055	1.593	H
HETATM	80	Zn	0	-0.173	-2.001	1.352	Zn
HETATM	81	C	0	4.158	-4.589	0.103	C
HETATM	82	O	0	4.743	-5.806	-0.323	O
HETATM	83	C	0	5.552	-5.720	-1.512	C
HETATM	84	C	0	6.916	-5.101	-1.182	C
HETATM	85	H	0	6.823	-4.035	-0.921	H
HETATM	86	H	0	7.559	-5.148	-2.067	H
HETATM	87	O	0	7.549	-5.832	-0.154	O
HETATM	88	H	0	6.874	-5.993	0.522	H
HETATM	89	H	0	5.742	-6.768	-1.766	H
HETATM	90	H	0	4.914	-3.836	0.378	H
HETATM	91	C	0	3.220	-4.030	-1.001	C
HETATM	92	N	0	2.489	-2.809	-0.515	N
HETATM	93	H	0	2.233	-3.008	0.471	H
HETATM	94	C	0	1.134	-2.362	-1.179	C
HETATM	95	O	0	0.075	-2.749	-0.433	O
HETATM	96	C	0	1.003	-2.876	-2.614	C
HETATM	97	H	0	1.090	-3.962	-2.667	H
HETATM	98	H	0	1.740	-2.412	-3.270	H
HETATM	99	H	0	0.011	-2.599	-2.972	H
HETATM	100	H	0	2.497	-4.819	-1.224	H
HETATM	101	C	0	4.056	-3.763	-2.267	C
HETATM	102	O	0	4.924	-2.651	-2.145	O
HETATM	103	H	0	5.369	-2.644	-1.287	H
HETATM	104	H	0	3.411	-3.528	-3.111	H
HETATM	105	C	0	4.787	-5.063	-2.668	C
HETATM	106	H	0	5.490	-4.808	-3.475	H
HETATM	107	O	0	3.751	-5.915	-3.150	O
HETATM	108	C	0	4.182	-6.930	-4.042	C
HETATM	109	H	0	3.289	-7.474	-4.358	H
HETATM	110	H	0	4.871	-7.643	-3.568	H
HETATM	111	H	0	4.675	-6.504	-4.928	H
HETATM	112	O	0	-0.661	-0.054	-3.045	O
HETATM	113	H	0	0.103	-0.304	-2.469	H
HETATM	114	H	0	-1.354	-0.711	-2.854	H
HETATM	115	O	0	1.184	-0.923	-1.198	O
HETATM	116	H	0	2.513	2.051	0.699	H
HETATM	117	H	0	2.103	-0.578	-1.079	H
HETATM	118	H	0	9.490	4.591	0.198	H
HETATM	119	H	0	-4.249	-6.145	3.857	H
HETATM	120	H	0	-4.989	1.877	2.358	H
HETATM	121	H	0	3.882	-1.372	3.418	H
HETATM	122	H	0	-1.245	3.306	1.499	H
HETATM	123	H	0	-4.058	5.328	-1.912	H
HETATM	124	H	0	3.933	2.032	3.022	H
HETATM	125	C	0	-4.666	0.245	-1.914	C
HETATM	126	C	0	-4.623	0.185	-3.435	C
HETATM	127	H	0	-3.615	0.434	-3.784	H
HETATM	128	H	0	-4.892	-0.797	-3.825	H
HETATM	129	H	0	-5.326	0.939	-3.845	H
HETATM	130	O	0	-4.889	1.329	-1.347	O
HETATM	131	N	0	-4.014	-3.620	-3.481	N
HETATM	132	H	0	-3.014	-3.487	-3.488	H
HETATM	133	C	0	-4.647	-4.473	-4.476	C
HETATM	134	H	0	-3.991	-4.623	-5.340	H
HETATM	135	H	0	-5.563	-3.982	-4.806	H
HETATM	136	O	0	3.252	-5.090	1.407	O
HETATM	137	C	0	2.833	-6.443	1.606	C
HETATM	138	C	0	4.205	-7.308	1.507	C
HETATM	139	H	0	2.407	-6.545	2.582	H
HETATM	140	C	0	1.305	-7.154	1.318	C
HETATM	141	C	0	3.755	-8.765	1.721	C
HETATM	142	H	0	4.632	-7.206	0.531	H
HETATM	143	O	0	5.182	-6.961	2.492	O
HETATM	144	H	0	1.404	-6.680	0.363	H
HETATM	145	O	0	1.310	-8.683	1.135	O
HETATM	146	C	0	0.165	-6.755	2.274	C
HETATM	147	H	0	3.328	-8.866	2.697	H
HETATM	148	C	0	2.703	-9.138	0.660	C
HETATM	149	N	0	4.914	-9.660	1.597	N
HETATM	150	H	0	5.299	-6.008	2.507	H
HETATM	151	H	0	0.537	-6.068	3.005	H
HETATM	152	H	0	-0.622	-6.291	1.717	H
HETATM	153	O	0	-0.338	-7.923	2.929	O
HETATM	154	H	0	2.702	-10.198	0.516	H
HETATM	155	H	0	2.941	-8.653	-0.264	H
HETATM	156	H	0	5.114	-10.070	2.487	H
HETATM	157	C	0	6.242	-9.459	0.712	C
HETATM	158	H	0	-0.596	-8.573	2.271	H
HETATM	159	C	0	7.521	-10.306	0.847	C
HETATM	160	O	0	6.142	-8.654	-0.251	O
HETATM	161	H	0	8.344	-9.671	1.104	H
HETATM	162	H	0	7.384	-11.040	1.612	H
HETATM	163	H	0	7.725	-10.794	-0.083	H
HETATM	164	C	0	-4.996	-5.846	-3.872	C
HETATM	165	H	0	-5.944	-5.789	-3.380	H
HETATM	166	H	0	-5.043	-6.577	-4.652	H
HETATM	167	C	0	-3.912	-6.249	-2.855	C
HETATM	168	C	0	-4.280	-6.681	-1.580	C
HETATM	169	C	0	-2.565	-6.185	-3.208	C
HETATM	170	C	0	-3.299	-7.047	-0.658	C
HETATM	171	H	0	-5.342	-6.732	-1.301	H
HETATM	172	C	0	-1.583	-6.551	-2.286	C
HETATM	173	H	0	-2.275	-5.846	-4.213	H
HETATM	174	C	0	-1.951	-6.982	-1.011	C
HETATM	175	H	0	-3.588	-7.387	0.347	H
HETATM	176	H	0	-0.521	-6.500	-2.565	H
HETATM	177	O	0	-0.945	-7.357	-0.066	O
HETATM	178	H	0	-0.738	-8.289	-0.171	H

EP

HETATM 1 C 0 -0.278 3.029 1.918 C
HETATM 2 H 0 -0.076 3.642 2.799 H
HETATM 3 H 0 0.064 2.012 2.140 H
HETATM 4 C 0 0.486 3.548 0.739 C
HETATM 5 O 0 0.048 3.663 -0.400 O
HETATM 6 O 0 1.744 3.859 1.058 O
HETATM 7 C 0 2.914 -0.681 3.524 C
HETATM 8 H 0 3.285 -1.502 4.137 H
HETATM 9 H 0 2.708 0.201 4.133 H
HETATM 10 C 0 1.691 -1.115 2.752 C
HETATM 11 O 0 0.834 -0.186 2.442 O
HETATM 12 O 0 1.543 -2.301 2.398 O
HETATM 13 C 0 -4.870 2.494 2.889 C
HETATM 14 H 0 -4.768 3.563 2.671 H
HETATM 15 H 0 -4.757 2.359 3.967 H
HETATM 16 C 0 -3.849 1.687 2.161 C
HETATM 17 N 0 -3.797 1.614 0.776 N
HETATM 18 H 0 -4.559 1.830 0.122 H
HETATM 19 C 0 -2.783 0.790 0.431 C
HETATM 20 H 0 -2.550 0.513 -0.582 H
HETATM 21 N 0 -2.158 0.347 1.509 N
HETATM 22 C 0 -2.811 0.903 2.600 C
HETATM 23 H 0 -2.476 0.716 3.609 H
HETATM 24 C 0 -4.275 -4.228 4.551 C
HETATM 25 H 0 -5.102 -3.517 4.483 H
HETATM 26 H 0 -4.065 -4.395 5.614 H
HETATM 27 C 0 -3.084 -3.690 3.832 C
HETATM 28 N 0 -1.875 -4.366 3.752 N
HETATM 29 H 0 -1.671 -5.261 4.173 H
HETATM 30 C 0 -0.998 -3.620 3.040 C
HETATM 31 H 0 0.034 -3.871 2.848 H
HETATM 32 N 0 -1.580 -2.497 2.648 N
HETATM 33 C 0 -2.876 -2.527 3.138 C
HETATM 34 H 0 -3.561 -1.711 2.965 H
HETATM 35 C 0 -3.412 4.464 -1.759 C
HETATM 36 H 0 -2.615 4.584 -1.020 H
HETATM 37 H 0 -4.149 3.738 -1.402 H
HETATM 38 N 0 -2.863 4.041 -3.050 N
HETATM 39 H 0 -3.238 4.479 -3.878 H
HETATM 40 C 0 -2.166 2.905 -3.228 C
HETATM 41 N 0 -1.599 2.297 -2.176 N
HETATM 42 H 0 -1.232 2.857 -1.397 H
HETATM 43 H 0 -1.064 1.453 -2.349 H
HETATM 44 N 0 -2.051 2.373 -4.453 N
HETATM 45 H 0 -2.690 2.653 -5.181 H
HETATM 46 H 0 -1.613 1.443 -4.521 H
HETATM 47 N 0 -4.939 -0.901 -1.120 N
HETATM 48 C 0 -4.071 -2.042 -1.467 C
HETATM 49 O 0 -2.790 -1.604 -1.886 O
HETATM 50 H 0 -2.129 -1.751 -1.169 H
HETATM 51 C 0 -4.717 -2.936 -2.551 C
HETATM 52 O 0 -5.918 -3.175 -2.501 O
HETATM 53 C 0 -4.035 -2.843 -0.146 C
HETATM 54 H 0 -3.231 -2.439 0.478 H
HETATM 55 H 0 -3.829 -3.901 -0.320 H
HETATM 56 C 0 -5.400 -2.562 0.487 C
HETATM 57 H 0 -5.414 -2.774 1.558 H
HETATM 58 H 0 -6.169 -3.156 -0.010 H
HETATM 59 C 0 -5.606 -1.074 0.184 C
HETATM 60 H 0 -5.136 -0.450 0.951 H
HETATM 61 H 0 -6.653 -0.776 0.099 H
HETATM 62 C 0 9.723 3.642 0.691 C
HETATM 63 H 0 9.893 4.542 1.281 H
HETATM 64 H 0 10.046 3.800 -0.342 H
HETATM 65 C 0 8.245 3.285 0.720 C
HETATM 66 O 0 7.980 2.187 0.007 O
HETATM 67 O 0 7.405 3.928 1.326 O
HETATM 68 C 0 4.655 2.202 2.221 C
HETATM 69 H 0 5.341 3.044 2.090 H
HETATM 70 H 0 5.160 1.482 2.877 H
HETATM 71 C 0 4.331 1.593 0.893 C
HETATM 72 N 0 5.297 1.406 -0.084 N
HETATM 73 H 0 6.244 1.721 -0.020 H
HETATM 74 C 0 4.688 0.829 -1.107 C
HETATM 75 H 0 5.157 0.503 -2.023 H
HETATM 76 N 0 3.366 0.644 -0.848 N
HETATM 77 H 0 3.654 -1.667 0.133 H
HETATM 78 C 0 3.125 1.119 0.427 C
HETATM 79 H 0 2.156 1.044 0.893 H
HETATM 80 Zn 0 -0.645 -1.044 1.506 Zn
HETATM 81 C 0 3.894 -4.463 -0.045 C
HETATM 82 O 0 4.210 -5.614 -0.828 O
HETATM 83 C 0 5.081 -5.365 -1.945 C
HETATM 84 C 0 6.531 -5.250 -1.458 C
HETATM 85 H 0 6.690 -4.322 -0.882 H
HETATM 86 H 0 7.206 -5.199 -2.319 H
HETATM 87 O 0 6.883 -6.381 -0.695 O
HETATM 88 H 0 6.112 -6.579 -0.140 H
HETATM 89 H 0 5.027 -6.294 -2.523 H
HETATM 90 H 0 4.785 -4.065 0.471 H
HETATM 91 C 0 3.234 -3.371 -0.890 C
HETATM 92 N 0 2.827 -2.198 -0.129 N
HETATM 93 H 0 2.361 -2.456 0.739 H
HETATM 94 C 0 0.076 -1.449 -1.363 C
HETATM 95 O 0 -0.596 -1.817 -0.307 O
HETATM 96 C 0 0.318 -2.556 -2.369 C
HETATM 97 H 0 0.368 -3.524 -1.875 H
HETATM 98 H 0 1.235 -2.370 -2.926 H
HETATM 99 H 0 -0.514 -2.573 -3.081 H
HETATM 100 H 0 2.359 -3.834 -1.359 H
HETATM 101 C 0 4.167 -2.961 -2.036 C
HETATM 102 O 0 5.276 -2.179 -1.598 O
HETATM 103 H 0 5.787 -2.667 -0.938 H
HETATM 104 H 0 3.625 -2.307 -2.725 H
HETATM 105 C 0 4.590 -4.208 -2.837 C
HETATM 106 H 0 5.401 -3.913 -3.523 H
HETATM 107 O 0 3.426 -4.583 -3.568 O
HETATM 108 C 0 3.678 -5.359 -4.727 C
HETATM 109 H 0 2.711 -5.534 -5.206 H
HETATM 110 H 0 4.128 -6.334 -4.494 H
HETATM 111 H 0 4.337 -4.830 -5.432 H
HETATM 112 O 0 -1.143 -0.188 -3.969 O
HETATM 113 H 0 -0.349 -0.210 -3.396 H
HETATM 114 H 0 -1.813 -0.663 -3.449 H
HETATM 115 O 0 0.349 -0.262 -1.635 O
HETATM 116 H 0 2.242 4.063 0.248 H
HETATM 117 H 0 1.427 -0.067 -1.534 H
HETATM 118 H 0 10.311 2.811 1.092 H
HETATM 119 H 0 -4.612 -5.178 4.122 H
HETATM 120 H 0 -5.887 2.195 2.613 H
HETATM 121 H 0 3.686 -0.413 2.796 H
HETATM 122 H 0 -1.343 3.018 1.697 H
HETATM 123 H 0 -3.893 5.432 -1.898 H
HETATM 124 H 0 3.745 2.540 2.722 H
HETATM 125 C 0 -5.102 0.231 -1.856 C
HETATM 126 C 0 -4.687 0.228 -3.313 C
HETATM 127 H 0 -3.653 0.569 -3.390 H
HETATM 128 H 0 -4.759 -0.752 -3.786 H
HETATM 129 H 0 -5.326 0.939 -3.845 H
HETATM 130 O 0 -5.568 1.266 -1.345 O
HETATM 131 N 0 -3.857 -3.446 -3.463 N
HETATM 132 H 0 -2.890 -3.167 -3.393 H
HETATM 133 C 0 -4.270 -4.388 -4.493 C
HETATM 134 H 0 -3.576 -4.354 -5.306 H
HETATM 135 H 0 -5.246 -4.126 -4.846 H
HETATM 136 O 0 2.999 -4.911 0.976 O
HETATM 137 C 0 3.275 -5.517 2.401 C
HETATM 138 C 0 4.686 -5.943 3.084 C
HETATM 139 C 0 1.825 -6.357 2.743 C
HETATM 140 C 0 4.277 -6.696 4.365 C
HETATM 141 H 0 5.221 -6.606 2.437 H
HETATM 142 O 0 5.521 -4.835 3.429 O
HETATM 143 H 0 2.016 -6.820 1.798 H
HETATM 144 O 0 1.920 -7.411 3.861 O
HETATM 145 C 0 0.661 -5.697 3.271 C
HETATM 146 H 0 3.742 -6.033 5.012 H
HETATM 147 C 0 3.378 -7.889 3.994 C
HETATM 148 N 0 5.482 -7.185 5.051 N
HETATM 149 H 0 5.593 -4.242 2.677 H
HETATM 150 H 0 1.017 -4.777 3.683 H
HETATM 151 H 0 -0.031 -5.491 2.481 H
HETATM 152 O 0 0.006 -6.455 4.292 O
HETATM 153 H 0 3.441 -8.635 4.759 H
HETATM 154 H 0 3.702 -8.305 3.063 H
HETATM 155 H 0 5.592 -6.698 5.917 H
HETATM 156 C 0 6.896 -7.606 4.410 C
HETATM 157 H 0 0.654 -6.988 4.759 H
HETATM 158 C 0 8.186 -7.847 5.216 C
HETATM 159 O 0 6.877 -7.903 3.187 O
HETATM 160 H 0 8.943 -7.164 4.890 H
HETATM 161 H 0 7.990 -7.693 6.256 H
HETATM 162 H 0 8.521 -8.851 5.061 H
HETATM 163 C 0 -4.302 -5.811 -3.904 C
HETATM 164 H 0 -5.238 -5.971 -3.410 H
HETATM 165 H 0 -4.188 -6.525 -4.692 H
HETATM 166 C 0 -3.153 -5.974 -2.892 C
HETATM 167 C 0 -3.410 -6.492 -1.623 C
HETATM 168 C 0 -1.855 -5.605 -3.245 C
HETATM 169 C 0 -2.369 -6.641 -0.707 C
HETATM 170 H 0 -4.433 -6.784 -1.345 H
HETATM 171 C 0 -0.814 -5.753 -2.328 C
HETATM 172 H 0 -1.652 -5.196 -4.246 H
HETATM 173 C 0 -1.071 -6.271 -1.059 C
HETATM 174 H 0 -2.572 -7.050 0.294 H
HETATM 175 H 0 0.209 -5.462 -2.607 H
HETATM 176 O 0 -0.004 -6.423 -0.119 O
HETATM 177 H 0 0.405 -7.284 -0.236 H
HETATM 178 H 0 -0.738 -8.289 -0.171 H

File Glycan.prepc

0 0 2

This is a remark line
molecule.res
MOL XYZ 0
CHANGE OMIT DU BEG
0.0000

1	DUMM	DU	M	999.000	999.0	-999.0	.00000
2	DUMM	DU	M	999.000	-999.0	999.0	.00000
3	DUMM	DU	M	-999.000	999.0	999.0	.00000
4	N1	n7	M	-0.829000	8.630000	-1.744000	-0.756081
5	H1	hn	E	-1.735000	8.454000	-2.132000	0.403466
6	C27	c3	3	-0.441000	10.019000	-1.889000	0.117146
7	H39	h1	E	-0.485000	10.284000	-2.938000	0.027289
8	H40	h1	E	-1.058000	10.726000	-1.337000	0.027289
9	H41	h1	E	0.585000	10.142000	-1.561000	0.027289
10	C1	c3	M	-0.744000	8.065000	-0.412000	0.164057
11	H2	h1	E	0.252000	8.270000	-0.034000	0.180464
12	C2	c3	3	-1.782000	8.610000	0.581000	-0.391288
13	H3	h1	E	-1.748000	9.689000	0.588000	0.187596
14	H4	h1	E	-2.767000	8.287000	0.275000	0.187596
15	S1	sh	S	-1.432000	7.996000	2.268000	-0.368776
16	H5	hs	E	-2.528000	8.459000	2.854000	0.245837
17	C3	c	M	-0.938000	6.562000	-0.593000	0.487073
18	O1	o	E	-1.986000	6.120000	-0.990000	-0.547012
19	N2	ns	M	0.120000	5.777000	-0.315000	-0.408734
20	H6	hn	E	0.992000	6.162000	-0.028000	0.290967
21	C4	c3	M	0.052000	4.346000	-0.447000	0.011140
22	H7	h1	E	-0.929000	4.020000	-0.133000	0.133072
23	C6	c	B	1.144000	3.758000	0.449000	0.726094
24	O2	o	E	2.202000	4.319000	0.556000	-0.583778
25	N3	ns	B	0.835000	2.606000	1.072000	-0.705547
26	H10	hn	E	-0.029000	2.161000	0.855000	0.274625
27	C8	c3	3	1.808000	1.826000	1.816000	0.297546
28	H11	h1	E	2.694000	2.435000	1.901000	0.055417
29	C9	c3	3	1.275000	1.484000	3.218000	-0.140355
30	H12	hc	E	1.033000	2.422000	3.706000	0.070412
31	H13	hc	E	0.348000	0.932000	3.111000	0.070412
32	C11	ca	S	2.253000	0.688000	4.056000	0.055763
33	C12	ca	B	2.083000	-0.671000	4.265000	-0.189731
34	H14	ha	E	1.236000	-1.175000	3.833000	0.167769
35	C14	ca	B	2.983000	-1.404000	5.026000	-0.309019
36	H16	ha	E	2.824000	-2.458000	5.181000	0.188281
37	C16	ca	B	4.077000	-0.774000	5.592000	0.427707
38	C15	ca	B	4.268000	0.588000	5.395000	-0.309019
39	C13	ca	S	3.362000	1.301000	4.639000	-0.189731
40	H15	ha	E	3.514000	2.360000	4.510000	0.167769
41	H17	ha	E	5.116000	1.067000	5.848000	0.188281
42	O6	oh	S	4.990000	-1.429000	6.341000	-0.560874
43	H18	ho	E	4.715000	-2.320000	6.508000	0.386961
44	C10	c	B	2.102000	0.556000	1.018000	0.642566
45	O5	o	E	1.198000	-0.180000	0.687000	-0.624653
46	N4	ns	B	3.383000	0.325000	0.708000	-0.609692
47	H19	hn	E	4.097000	0.967000	0.973000	0.338653
48	C17	c3	3	3.780000	-0.778000	-0.141000	-0.017583
49	H20	h1	E	3.126000	-1.611000	0.076000	0.098125
50	C18	c3	3	3.652000	-0.375000	-1.623000	-0.107309
51	H21	hc	E	2.735000	0.189000	-1.709000	0.072499
52	H22	hc	E	4.465000	0.302000	-1.866000	0.072499
53	C20	c3	3	3.611000	-1.539000	-2.615000	0.092141
54	H23	hc	E	4.548000	-2.084000	-2.628000	-0.018307
55	H24	hc	E	2.834000	-2.239000	-2.307000	-0.018307
56	C21	c3	3	3.280000	-1.018000	-4.019000	0.007585
57	H25	hc	E	2.420000	-0.357000	-3.959000	0.017353
58	H26	hc	E	4.105000	-0.420000	-4.394000	0.017353
59	C22	c3	3	2.972000	-2.139000	-5.005000	0.035196
60	H27	hx	E	3.862000	-2.655000	-5.336000	0.077778
61	H28	hx	E	2.283000	-2.861000	-4.593000	0.077778
62	N5	nz	3	2.286000	-1.572000	-6.217000	-0.370399
63	H29	hn	E	2.826000	-0.831000	-6.638000	0.320620
64	H30	hn	E	2.128000	-2.272000	-6.925000	0.320620
65	H31	hn	E	1.364000	-1.189000	-5.945000	0.320620
66	C19	c	B	5.227000	-1.116000	0.195000	0.577343
67	O7	o	E	6.023000	-0.237000	0.395000	-0.547964
68	N6	ns	B	5.538000	-2.425000	0.231000	-0.276792
69	H32	hn	E	4.857000	-3.102000	-0.031000	0.264077
70	C23	c3	3	6.897000	-2.887000	0.310000	-0.337401
71	H33	h1	E	7.488000	-2.122000	0.791000	0.163897
72	C24	c3	3	6.980000	-4.179000	1.129000	0.419948
73	H34	h1	E	8.002000	-4.536000	1.082000	-0.013231
74	C26	c3	3	6.581000	-3.982000	2.585000	-0.183610
75	H35	hc	E	5.567000	-3.611000	2.666000	0.046326
76	H36	hc	E	6.644000	-4.927000	3.118000	0.046326
77	H37	hc	E	7.247000	-3.281000	3.077000	0.046326
78	O8	oh	S	6.124000	-5.089000	0.471000	-0.716792
79	H38	ho	E	6.164000	-5.937000	0.891000	0.452173
80	C25	c	B	7.476000	-3.088000	-1.089000	0.896195
81	O9	o	E	6.877000	-2.925000	-2.104000	-0.596966
82	O10	os	S	8.742000	-3.443000	-1.030000	-0.404778
83	C28	c3	3	9.428000	-3.647000	-2.257000	0.021216
84	H42	h1	E	10.434000	-3.923000	-1.988000	0.076972
85	H43	h1	E	8.957000	-4.438000	-2.822000	0.076972
86	H44	h1	E	9.430000	-2.737000	-2.839000	0.076972
87	C5	c3	M	0.274000	3.925000	-1.924000	-0.441195
88	H8	hc	E	-0.420000	4.489000	-2.529000	0.162939
89	H9	hc	E	1.292000	4.148000	-2.214000	0.162939
90	C7	c	M	0.025000	2.452000	-2.106000	0.860251
91	O4	o	E	0.865000	1.615000	-2.050000	-0.574723
92	O3	os	M	-1.271000	2.189000	-2.326000	-0.503686
93	C29	c3	M	-1.934000	1.062000	-1.825000	0.364425
94	O11	os	S	-2.089000	1.274000	-0.456000	-0.404417
95	C31	c3	B	-2.340000	0.158000	0.406000	0.171337
96	C34	c3	3	-3.092000	0.721000	1.596000	0.150747
97	H45	h1	E	-3.210000	-0.038000	2.352000	0.054400
98	H46	h1	E	-2.509000	1.529000	2.027000	0.054400
99	O12	oh	S	-4.375000	1.168000	1.245000	-0.646086
100	H47	ho	E	-4.305000	1.990000	0.778000	0.439443
101	H65	h1	E	-1.395000	-0.244000	0.751000	0.092329
102	H61	h2	E	-2.894000	1.109000	-2.313000	0.129034
103	C30	c3	M	-1.275000	-0.306000	-2.018000	-0.043651
104	N7	ns	B	-0.641000	-0.577000	-3.296000	-0.587124
105	H48	hn	E	0.300000	-0.882000	-3.197000	0.360560
106	C35	c	B	-0.974000	-0.387000	-4.578000	0.914757
107	O13	o	E	-0.170000	-0.696000	-5.452000	-0.763523
108	C36	c3	3	-2.282000	0.247000	-4.961000	-0.532446
109	H49	hc	E	-2.209000	1.318000	-4.806000	0.167530
110	H50	hc	E	-2.448000	0.062000	-6.012000	0.167530
111	H51	hc	E	-3.102000	-0.143000	-4.380000	0.167530
112	H62	h1	E	-0.459000	-0.334000	-1.316000	0.196056
113	C32	c3	M	-2.283000	-1.388000	-1.592000	0.095890
114	O14	oh	S	-3.108000	-1.682000	-2.679000	-0.697202
115	H52	ho	E	-3.774000	-2.291000	-2.375000	0.482030
116	H63	h1	E	-1.724000	-2.269000	-1.292000	0.116481
117	C33	c3	M	-3.103000	-0.915000	-0.371000	-0.031066

```
 118  H64   h1    E         -4.028000   -0.485000   -0.729000    0.111364
 119  O15   os    M         -3.346000   -1.988000    0.505000   -0.415661
 120  C37   c3    M         -4.523000   -2.716000    0.439000    0.281689
 121  O16   os    S         -4.790000   -3.051000   -0.896000   -0.300729
 122  C39   c3    B         -5.833000   -3.988000   -1.138000   -0.052282
 123  C42   c3    3         -5.233000   -5.316000   -1.557000    0.211474
 124  H53   h1    E         -6.038000   -5.992000   -1.825000    0.038452
 125  H54   h1    E         -4.608000   -5.160000   -2.432000    0.038452
 126  O17   oh    S         -4.474000   -5.818000   -0.489000   -0.712734
 127  H55   ho    E         -4.162000   -6.689000   -0.692000    0.452266
 128  H69   h1    E         -6.406000   -3.595000   -1.968000    0.080637
 129  H70   h2    E         -4.316000   -3.610000    1.001000    0.132126
 130  C38   c3    M         -5.750000   -1.959000    0.961000   -0.013376
 131  N8    ns    B         -5.622000   -1.413000    2.298000   -0.606986
 132  H57   hn    E         -5.535000   -0.418000    2.312000    0.329572
 133  C43   c     B         -5.071000   -1.958000    3.433000    0.776912
 134  O20   o     E         -4.683000   -1.230000    4.304000   -0.622469
 135  C44   c3    3         -4.974000   -3.459000    3.585000   -0.310019
 136  H58   hc    E         -3.977000   -3.779000    3.298000    0.088586
 137  H59   hc    E         -5.099000   -3.686000    4.634000    0.088586
 138  H60   hc    E         -5.709000   -3.997000    3.009000    0.088586
 139  H66   h1    E         -5.831000   -1.085000    0.331000    0.110413
 140  C40   c3    M         -7.060000   -2.788000    0.730000    0.237934
 141  O19   oh    S         -7.733000   -3.070000    1.920000   -0.622362
 142  H56   ho    E         -7.814000   -2.275000    2.431000    0.416748
 143  H67   h1    E         -7.710000   -2.218000    0.072000    0.042508
 144  C41   c3    M         -6.791000   -4.130000    0.050000    0.177601
 145  H68   h1    E         -6.367000   -4.806000    0.784000    0.116168
 146  O18   os    M         -7.951000   -4.683000   -0.494000   -0.419785
 147  C45   c3    M         -8.719000   -5.519000    0.335000    0.118148
 148  H71   h1    E         -9.161000   -4.971000    1.154000    0.037656
 149  H72   h1    E         -9.498000   -5.937000   -0.287000    0.037656
 150  H73   h1    E         -8.113000   -6.329000    0.733000    0.037656


LOOP
  C13  C11
  C33  C31
  C41  C39

IMPROPER
   C1   N2   C3   O1
   C3   C4   N2   H6
   C4   N3   C6   O2
   C6   C8   N3  H10
   C9  C12  C11  C13
  C11  C14  C12  H14
  C12  C16  C14  H16
  C14  C15  C16   O6
  C16  C13  C15  H17
  C11  C15  C13  H15
   C8   N4  C10   O5
  C17  C10   N4  H19
  C17   N6  C19   O7
  C23  C19   N6  H32
  C23   O9  C25  O10
   C5   O4   C7   O3
  C35  C30   N7  H48
  C36   N7  C35  O13
  C43  C38   N8  H57
  C44   N8  C43  O20

DONE
STOP


File Glycan.frcmod

remark goes here
MASS
n7 14.010        0.530
hn 1.008         0.161
c3 12.010        0.878
h1 1.008         0.135
sh 32.060        2.900
hs 1.008         0.135
c  12.010        0.616
o  16.000        0.434
ns 14.010        0.530
hc 1.008         0.135
ca 12.010        0.360
ha 1.008         0.135
oh 16.000        0.465
ho 1.008         0.135
hx 1.008         0.135
nz 14.010        0.530
os 16.000        0.465
h2 1.008         0.135

BOND
n7-hn  511.28   1.019       same as hn-n3
n7-c3  261.19   1.465       same as c3-n3
c3-h1  375.92   1.097
c3-c3  232.52   1.538
c3-c   243.22   1.524
c3-sh  180.78   1.843
sh-hs  294.59   1.347
c -o   652.57   1.218
c -ns  356.21   1.379       same as c -n
ns-hn  527.31   1.013       same as hn-n
ns-c3  263.77   1.462       same as c3-n
c3-hc  375.92   1.097
c3-ca  250.32   1.516
ca-ca  378.57   1.398
ca-ha  395.72   1.086
ca-oh  365.55   1.364
oh-ho  563.51   0.973
c3-hx  386.49   1.091
c3-nz  222.58   1.511       same as c3-n4
nz-hn  482.87   1.030       same as hn-n4
c3-oh  293.40   1.423
c -os  372.94   1.358
os-c3  284.76   1.432
c3-h2  377.33   1.096

ANGLE
n7-c3-h1   61.163     109.880   same as h1-c3-n3
n7-c3-c3   83.305     111.040   same as c3-c3-n3
n7-c3-c    83.673     111.140   same as c -c3-n3
hn-n7-c3   47.782     109.290   same as c3-n3-hn
c3-n7-c3   65.697     112.350   same as c3-n3-c3
h1-c3-h1   38.802     108.460
c3-c3-h1   46.868     109.560
c3-c3-sh   62.313     113.130
c3-c -o    84.552     123.200
c3-c -ns   84.266     115.180   same as c3-c -n
```

```
h1-c3-c    47.531     108.220
c3-c3-c    65.307     111.040
c3-sh-hs   51.361      96.400
h1-c3-sh   42.420     108.420
c -ns-hn   48.691     117.550   same as c -n -hn
c -ns-c3   65.252     120.690   same as c -n -c3
o -c -ns  113.811     123.050   same as n -c -o
ns-c3-h1   61.544     108.880   same as h1-c3-n
ns-c3-c    84.540     109.060   same as c -c3-n
ns-c3-c3   83.161     111.610   same as c3-c3-n
hn-ns-c3   46.147     117.680   same as c3-n -hn
c3-c3-hc   46.816     109.800
c3-c3-ca   65.183     112.070
c3-ca-ca   65.583     120.770
hc-c3-hc   38.960     107.580
hc-c3-ca   47.281     110.470
ca-ca-ha   48.680     119.880
ca-ca-ca   68.767     120.020
ca-ca-oh   87.211     119.900
ca-oh-ho   50.712     108.580
c3-c3-c3   64.888     111.510
c3-c3-hx   46.677     110.560
c3-c3-nz   80.976     114.210   same as c3-c3-n4
c3-nz-hn   46.193     110.110   same as c3-n4-hn
hx-c3-hx   38.782     109.750
hx-c3-nz   60.076     108.010   same as hx-c3-n4
hn-nz-hn   40.020     108.300   same as hn-n4-hn
c3-c3-oh   84.642     110.190
c3-c -os   86.419     110.720
c3-oh-ho   49.027     107.260
h1-c3-oh   62.540     110.260
c -os-c3   66.906     115.980
o -c -os  114.822     123.250
os-c3-h1   62.377     109.780
hc-c3-c    47.411     108.770
os-c3-os  110.893     108.290
os-c3-h2   62.442     109.580
os-c3-c3   85.306     107.970
c3-os-c3   66.293     112.480
h2-c3-c3   46.730     110.220

DIHE
n7-c3-c3-h1   1    0.156         0.000          3.000
n7-c3-c3-sh   1    0.156         0.000          3.000
n7-c3-c -o    1    0.000       180.000          2.000
n7-c3-c -ns   1    0.000       180.000          2.000
hn-n7-c3-h1   1    0.300         0.000          3.000      same as X -c3-n3-X
hn-n7-c3-c3   1    0.217         0.000          3.000      same as hn-n3-c3-c3
hn-n7-c3-c    1    0.300         0.000          3.000      same as X -c3-n3-X
c3-n7-c3-h1   1    0.225         0.000          3.000      same as h1-c3-n3-c3
c3-n7-c3-c3   1    0.020       180.000         -3.000
c3-n7-c3-c3   1    0.050         0.000          2.000
c3-n7-c3-c    1    0.300         0.000          3.000      same as X -c3-n3-X
c3-c3-sh-hs   1    0.250         0.000          3.000
c3-c -ns-hn   1    2.500       180.000          2.000      same as X -c -n -X
c3-c -ns-c3   1    0.260       180.000         -2.000      same as c3-n -c -c3
c3-c -ns-c3   1    0.500         0.000          1.000      same as c3-n -c -c3
h1-c3-c3-h1   1    0.156         0.000          3.000
h1-c3-c3-sh   1    0.156         0.000          3.000
h1-c3-c -o    1    0.800         0.000         -1.000
h1-c3-c -o    1    0.080       180.000          3.000
h1-c3-c -ns   1    0.000       180.000          2.000
c3-c3-c -o    1    0.270       180.000          2.000
c3-c3-c -ns   1    0.000       180.000          2.000
h1-c3-c3-c    1    0.156         0.000          3.000
h1-c3-sh-hs   1    0.143         0.000          3.000
sh-c3-c3-c    1    0.156         0.000          3.000
c -ns-c3-h1   1    0.000       180.000          2.000      same as h1-c3-n -c
c -ns-c3-c    1    0.390       180.000         -2.000      same as c -c3-n -c
c -ns-c3-c    1    0.640         0.000          1.000      same as c -c3-n -c
c -ns-c3-c3   1    0.100       180.000         -4.000      same as c -n -c3-c3
c -ns-c3-c3   1    0.170         0.000         -3.000      same as c -n -c3-c3
c -ns-c3-c3   1    1.020       180.000          1.000      same as c -n -c3-c3
o -c -ns-hn   1    2.500       180.000         -2.000      same as hn-n -c -o
o -c -ns-hn   1    2.000         0.000          1.000      same as hn-n -c -o
o -c -ns-c3   1    2.500       180.000          2.000      same as X -c -n -X
ns-c3-c -o    1    0.000       180.000          2.000
ns-c3-c -ns   1    0.000       180.000          2.000
ns-c3-c3-hc   1    0.156         0.000          3.000
ns-c3-c3-c    1    0.156         0.000          3.000
hn-ns-c3-h1   1    0.000         0.000          2.000      same as X -c3-n -X
hn-ns-c3-c    1    0.000         0.000          2.000      same as X -c3-n -X
hn-ns-c3-c3   1    0.000         0.000          2.000      same as X -c3-n -X
c3-c3-c -os   1    0.000       180.000          2.000
h1-c3-c3-hc   1    0.156         0.000          3.000
c -c3-c3-hc   1    0.156         0.000          3.000
c -c3-c3-c    1    0.156         0.000          3.000
ns-c3-c3-ca   1    0.156         0.000          3.000
c3-c3-ca-ca   1    0.245       180.000          2.000
h1-c3-c3-ca   1    0.156         0.000          3.000
c3-ca-ca-ha   1    3.625       180.000          2.000
c3-ca-ca-ca   1    3.625       180.000          2.000
hc-c3-ca-ca   1    0.000       180.000          2.000
ca-c3-c3-c    1    0.100         0.000          3.000
ca-ca-ca-ha   1    3.625       180.000          2.000
ca-ca-ca-ca   1    3.625       180.000          2.000
ca-ca-ca-oh   1    3.625       180.000          2.000
ha-ca-ca-ha   1    3.625       180.000          2.000
ca-ca-oh-ho   1    0.835       180.000          2.000
ha-ca-ca-oh   1    3.625       180.000          2.000
ns-c3-c3-c3   1    0.156         0.000          3.000
c3-c3-c3-hc   1    0.080         0.000          3.000
c3-c3-c3-c3   1    0.130         0.000         -3.000
c3-c3-c3-c3   1    0.290       180.000         -2.000
c3-c3-c3-c3   1    0.110         0.000          1.000
h1-c3-c3-c3   1    0.156         0.000          3.000
hc-c3-c3-hc   1    0.120         0.000          3.000
c3-c3-c3-c    1    0.100         0.000          3.000
c3-c3-c3-hx   1    0.156         0.000          3.000
c3-c3-c3-nz   1    0.156         0.000          3.000
c3-c3-nz-hn   1    0.156         0.000          3.000      same as X -c3-n4-X
hc-c3-c3-hx   1    0.156         0.000          3.000
hc-c3-c3-nz   1    0.156         0.000          3.000
hx-c3-nz-hn   1    0.109         0.000          3.000      same as hn-n4-c3-hx
ns-c3-c3-h1   1    0.156         0.000          3.000
ns-c3-c3-oh   1    0.156         0.000          3.000
ns-c3-c -os   1    0.000       180.000          2.000
c3-c3-oh-ho   1    0.000         0.000          3.000
c3-c -os-c3   1    1.580       180.000         -1.000
c3-c -os-c3   1    3.180       180.000         -2.000
c3-c -os-c3   1    0.730         0.000          3.000
h1-c3-c3-oh   1    0.000         0.000         -3.000
h1-c3-c3-oh   1    0.250         0.000          1.000
h1-c3-c -os   1    0.000       180.000          2.000
h1-c3-oh-ho   1    0.113         0.000          3.000
hc-c3-c3-oh   1    0.180         0.000         -3.000
```

```
hc-c3-c3-oh   1    0.510         0.000          1.000
oh-c3-c3-c    1    0.210         0.000          3.000
c -os-c3-h1   1    0.383         0.000          3.000
o -c -os-c3   1    2.700       180.000         -2.000
o -c -os-c3   1    1.400       180.000          1.000
hc-c3-c -o    1    0.830         0.000         -1.000
hc-c3-c -o    1    0.040       180.000          3.000
hc-c3-c -os   1    0.000       180.000          2.000
c -os-c3-os   1    0.383         0.000          3.000
c -os-c3-h2   1    0.383         0.000          3.000
c -os-c3-c3   1    0.383         0.000         -3.000
c -os-c3-c3   1    0.800       180.000          1.000
os-c3-os-c3   1    0.000       180.000         -3.000
os-c3-os-c3   1    1.240         0.000         -2.000
os-c3-os-c3   1    0.970       180.000          1.000
os-c3-c3-ns   1    0.156         0.000          3.000
os-c3-c3-h1   1    0.000         0.000         -3.000
os-c3-c3-h1   1    0.250         0.000          1.000
os-c3-c3-c3   1    0.156         0.000          3.000
c3-os-c3-c3   1    0.240         0.000         -3.000
c3-os-c3-c3   1    0.160         0.000          2.000
c3-os-c3-h1   1    0.337         0.000          3.000
c3-c3-c3-oh   1    0.210         0.000          3.000
os-c3-c3-oh   1    1.010         0.000         -3.000
os-c3-c3-oh   1    0.000         0.000         -2.000
os-c3-c3-oh   1    0.020       180.000          1.000
os-c3-c3-os   1    0.000         0.000         -3.000
os-c3-c3-os   1    0.000       180.000         -2.000
os-c3-c3-os   1    0.170       180.000          1.000
c3-os-c3-h2   1    0.383         0.000          3.000
h2-c3-c3-ns   1    0.156         0.000          3.000
h2-c3-c3-h1   1    0.156         0.000          3.000
h2-c3-c3-c3   1    0.156         0.000          3.000
ns-c -c3-hc   1    0.000       180.000          2.000

IMPROPER
c3-ns-c -o         10.5          180.0        2.0          General improper torsional angle (2 general atom types)
c -c3-ns-hn         1.1          180.0        2.0          Using default value
c3-ca-ca-ca         1.1          180.0        2.0
ca-ca-ca-ha         1.1          180.0        2.0          General improper torsional angle (2 general atom types)
ca-ca-ca-oh         1.1          180.0        2.0          Using default value
c3-o -c -os        10.5          180.0        2.0          General improper torsional angle (2 general atom types)

NONBON
  n7          1.9686  0.0522
  hn          0.6210  0.0100
  c3          1.9069  0.1078
  h1          1.3593  0.0208
  sh          1.9825  0.2824
  hs          0.6112  0.0124
  c           1.8606  0.0988
  o           1.7107  0.1463
  ns          1.8352  0.1174
  hc          1.4593  0.0208
  ca          1.8606  0.0988
  ha          1.4735  0.0161
  oh          1.8200  0.0930
  ho          0.3019  0.0047
  hx          1.0593  0.0208
  nz          1.5528  1.1450
  os          1.7713  0.0726
  h2          1.2593  0.0208
```

Paper IV

"QM Cluster or QM/MM in Computational Enzymology: The Test Case of LigW-Decarboxylase"

Mario Prejanò, Tiziana Marino, Nino Russo
Frontiers in chemistry **2018**, *6***,** 249-257.

doi: 10.3389/fchem.2018.00249

frontiers in Chemistry

ORIGINAL RESEARCH
published: 28 June 2018
doi: 10.3389/fchem.2018.00249

QM Cluster or QM/MM in Computational Enzymology: The Test Case of LigW-Decarboxylase

Mario Prejanò, Tiziana Marino and Nino Russo*

Dipartimento Di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy

The catalytic mechanism of the decarboxylation of 5-carboxyvanillate by LigW producing vanillic acid has been studied by using QM cluster and hybrid QM/MM methodologies. In the QM cluster model, the environment of a small QM model is treated with a bulky potential while two QM/MM models studies include partial and full protein with and without explicitly treated water solvent. The studied reaction involves two sequential steps: the protonation of the carbon of the 5-carboxy-vanillate substrate and the decarboxylation of the intermediate from which results deprotonated vanillic acid as product. The structures and energetics obtained by using three structural models and two density functionals are quite consistent to each other. This indicates that the small QM cluster model of the presently considered enzymatic reaction is appropriate enough and the reaction is mainly influenced by the active site.

Keywords: QM, QM/MM, decarboxylation, enzymatic catalysis, reaction mechanism, LigW

OPEN ACCESS

Edited by:
Sam P. De Visser,
University of Manchester,
United Kingdom

Reviewed by:
Artur Nenov,
Università degli Studi di Bologna, Italy
Ahmet Altun,
Max-Planck-Institut für Kohlenforschung, Germany

****Correspondence:***
Tiziana Marino
tiziana.marino65@unical.it

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *11 April 2018*
Accepted: *08 June 2018*
Published: *28 June 2018*

Citation:
Prejanò M, Marino T and Russo N (2018) QM Cluster or QM/MM in Computational Enzymology: The Test Case of LigW-Decarboxylase. Front. Chem. 6:249. doi: 10.3389/fchem.2018.00249

INTRODUCTION

Enzymes are biological machines that efficiently catalyze a huge number of chemical reactions in the very short time steps required by the physiological processes. In the last decades, computational enzymology has become a very useful tool for studying enzyme activity since it allows to determine the energies and structures of short-lived intermediates and transition states. Through computational enzymology, different reaction pathways can be analyzed, and their feasibility can be established by a careful analysis of calculated energy barriers. A crucial issue in computational enzymology is the choice of the model to be used in the simulations. The choice is not so obvious because it depends on the nature of enzyme (without or with metal cofactor) on the catalytic pocket and on the amino acids implicated in the chemical reaction. In fully quantum mechanical (QM) treatment (Himo, 2006; Ramos and Fernandes, 2008; Siegbahn and Himo, 2011; Merz, 2014), a cluster that contains all the residues around the active site is considered.

In metalloenzymes, the construction of the cluster model is facilitated by the presence of metal ions and all the residues of their inner coordination sphere. One may also need to include some other surrounding residues involved in the chemical process. Atoms at the periphery of the model, where truncation is made, are normally frozen in their original positions present in the crystallographic structure for avoiding artificial expansion or other rearrangements (Blomberg et al., 2014). The surrounding protein environment not directly implicated in the chemical transformation, is modeled with implicit dielectric constant-based solvation models (Warshel, 1991). This method is highly versatile and widely applied to a large variety of enzyme families and to different classes of enzymes (Ramos and Fernandes, 2008; Liao et al., 2010; Amata et al., 2011a; Himo, 2017; Piazzetta et al., 2017; Prejanò et al., 2017a). A different approach developed in 1976

(Warshel and Levitt, 1976) is the hybrid quantum mechanics/molecular mechanics (QM/MM) (Senn and Thiel, 2009; Quesne et al., 2016; Ryde, 2016). In this procedure, other than the QM portion a large number of residues (or the whole enzyme sequence) is treated at molecular mechanics level (MM) (Senn and Thiel, 2009). Convergence studies performed by different research groups indicated that QM-cluster models (Siegbahn and Himo, 2011; Ryde, 2017) gives reliable energetics when the size of the model is large enough. Herein we perform a theoretical study using both QM cluster and QM/MM approaches on the gene product of LigW of 5-carboxyvanillate decarboxylase (5CVA) (Vladimirova et al., 2016). The QM part in all the models has been treated in the framework of density functional theory (DFT) and by using two different exchange-correlation functionals.

The LigW belongs to the amidohydrolase (AHS) superfamily including a high number of enzymes catalyzing the hydrolysis of a wide range of substrates. In all AHS members, a mononuclear or binuclear metal binding site is found (Gerlt and Babbitt, 2001; Seibert and Raushel, 2005). All AHS members have a $(\beta/\alpha)_8$-barrel structural fold and catalyze the metal-dependent hydrolysis of phosphate and carboxylate esters (Jackson et al., 2005; Shapir et al., 2006; Elias et al., 2008; Khurana et al., 2009; Duarte et al., 2011; Tobimatsu et al., 2013). LigW catalyzes the C-C bond cleavage of 5-CV to vanillate (VAN) in an oxidant-independent fashion. The 5-carboxyvanillate (5-CV) represents one of the final product of the multienzymatic degradation of the biphenyl lignin derivatives. The lignin degradation of microbial origin represents an interesting process from both commercial and biotechnological point of view owing to the plant biomass conversion in renewable aromatic chemicals and biofuels (Liu and Zhang, 2006). Furthermore, decarboxylation represents a process of widespread occurrence in nature and therefore it is of relevant biological interest (Faponle et al., 2016).

COMPUTATIONAL METHODS

All the calculations were carried out by using the Gaussian 09 program (Gaussian 09, Revision D.01, 2011)[1]. The QM portions were treated with the B3LYP (Lee et al., 1988; Becke, 1993) hybrid density functional. 6-31+G(d,p) basis set was used for the C, N, O, and H atoms, whereas the SDD pseudo-potential and corresponding orbital basis set (Andrae et al., 1990) were employed for Mn atom. Our own N-layered integrated molecular orbital and molecular mechanics (ONIOM) method was applied as the QM/MM method in the framework of electronic embedding scheme, in which the effects of the fixed MM charges are incorporated in the QM hamiltonian (Svensson et al., 1996; Vreven et al., 2006). As shown in Figure S1, the enzyme-substrate complex (ES) is a high-spin sextet species while its low-spin doublet and the intermediate-spin quartet states are energetically not accessible. The sextet state does not suffer from any spin contamination ($<S^2>$ equal to 8.75). The optimized minima and transition states on the potential energy surfaces were confirmed by the analysis of the corresponding Hessian matrices. Zero-point-energy corrections were calculated and added to the final energies. In order to obtain more accurate energies, single point calculations on the optimized structures were performed with the larger basis set 6-311+G(2d,2p) taking into account the effects of the protein environment by using the solvation model density (SMD) (Marenich et al., 2009), with a dielectric constant ($\varepsilon = 4$) of the enzyme environment, for the cluster simulations (Alberto et al., 2010; Liao et al., 2010; Amata et al., 2011a,b; Himo, 2017; Piazzetta et al., 2017; Prejanò et al., 2017a,b). Energetics presented includes D3 dispersion correction (Grimme et al., 2011). To evaluate the effect of the exchange-correlation functionals single point calculations on the B3LYP optimized geometries have been performed by using the M06-L functional that was previously demonstrated to be accurate for describing metal containing systems properties (Zhao and Truhlar, 2006, 2008) (see Table S1). NBO analysis (NBO, version 3.1, 2001)[2] was performed on all intercepted stationary points at QM and QM/MM levels with B3LYP functional. Furthermore, the noncovalent interactions on the minima of the PES have been assessed by using the NCIPLOT tool (NCIPLOT, version 3.0, 2011)[3].

COMPUTATIONAL SETUP AND QM MODEL DEFINITIONS

The model of the LigW active site, used for both QM and QM/MM calculations, was obtained from the three-dimensional structure of wild-type LigW in the presence of the substrate-like inhibitor 5-nitrovanillate (5-NV) isolated by *N. aromaticivorans* (PDB id: 4QRN, resolution: 1.07 Å). Vladimirova et al. (2016) due to the very small difference (one atom) between the inhibitor (5-NV) and substrate (5-CV). This choice has been already shown sufficient when structurally compared with larger QM clusters (Sheng et al., 2017). In the active site, (see **Figure 1**) the manganese ion is octahedrally coordinated to Glu-19, His-188, Asp-314, one water molecule **w1** and the substrate. Two water molecules, (**w2** and **w3**), located at about 5 A from the substrate and other residues of the active site pocket not directly bound to the metal ion are retained in QM region (Arg58, Phe212, His241, Arg252, and Tyr317).

In the QM/MM models, the Amber ff14SB force field (Maier et al., 2015) as implemented in AMBER16 software was used. The missing MM parameters for the substrate 5-CV were created from single molecule optimization at HF/6-31G(d) level of theory with the Antechamber tool, as implemented in AMBER16 (AMBER version 16, 2016)[4]. At this purpose the General Amber Force Field (GAFF) (Wang et al., 2004) and the Restrained Electrostatic Potential (RESP) (Bayly et al., 1993) methods were used to derive intramolecular and Lennard-Jones parameters and atomic charges, respectively (see Table S2).

[1]Gaussian 09, Revision D.01 (2011), *Gaussian*, Inc., Wallingford CT.

[2]NBO, version 3.1 (2001).

[3]NCIPLOT, version3.0 (2011). Download: http://www.lct.jussieu.fr/pagesperso/contrera/nciplot.html

[4]AMBER 16 (2016), University of California, San Francisco.

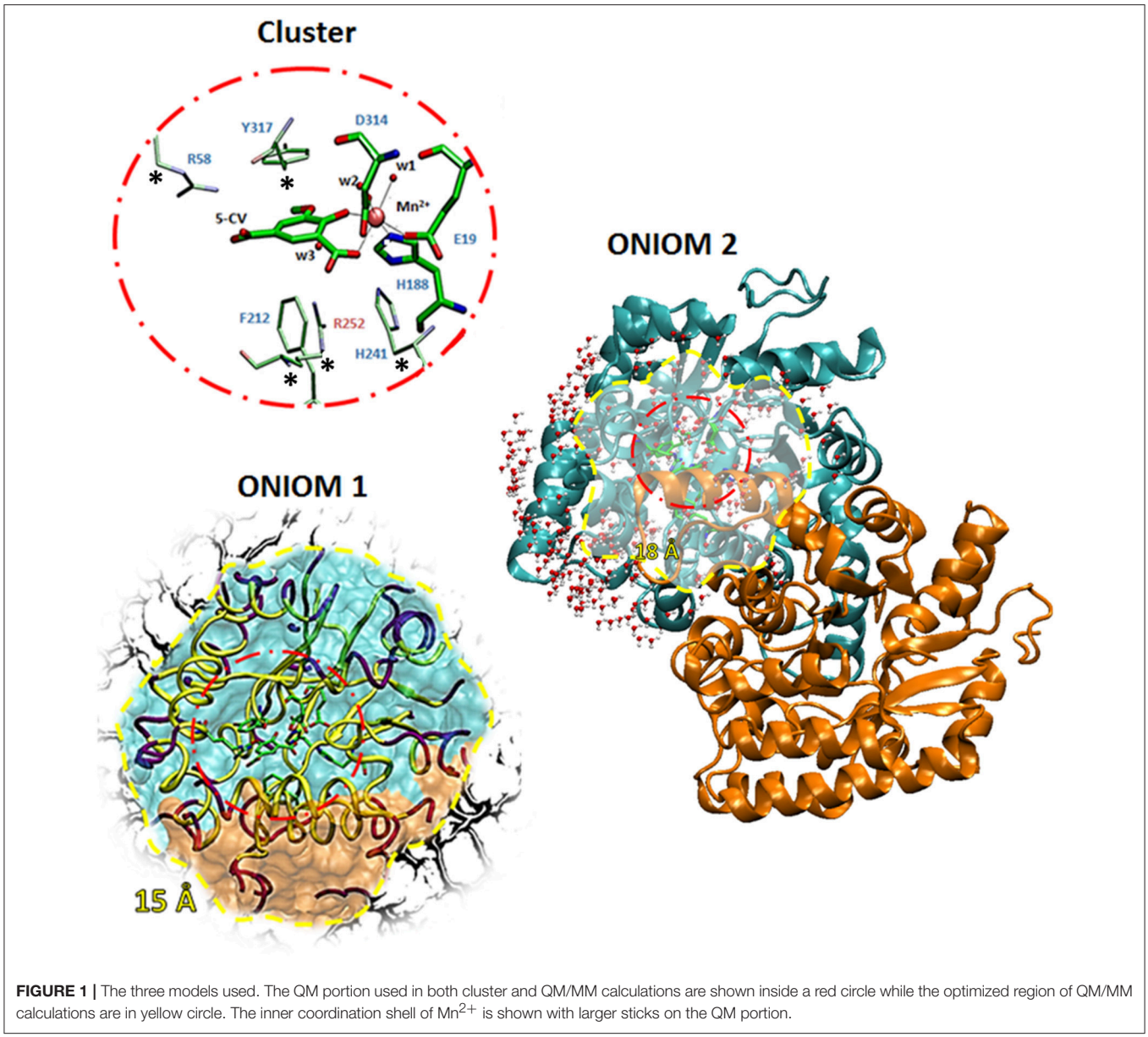


FIGURE 1 | The three models used. The QM portion used in both cluster and QM/MM calculations are shown inside a red circle while the optimized region of QM/MM calculations are in yellow circle. The inner coordination shell of Mn^{2+} is shown with larger sticks on the QM portion.

QM Cluster

All the amino acids of the QM region were truncated at the α-carbons, and hydrogen atoms were added manually. In order to avoid unrealistic movements of the groups during the geometry optimizations, the truncated α-carbons of the outer coordination shell labeled by stars in **Figure 1** were kept fixed to their crystallographic positions. The residues were modeled according to standard procedure (Liao et al., 2010; Amata et al., 2011a,b; Siegbahn and Himo, 2011; Blomberg et al., 2014; Himo, 2017; Prejanò et al., 2017a) considering the protonation states coming from the experimental evidences (Vladimirova et al., 2016). The obtained model consists of 126 atoms with a total charge equal to zero. The size of the cluster is adequate enough to represent the chemistry involved in the considered reaction mechanisms for formation or breaking bonds.

ONIOM-1

In this model, the QM region is surrounded by the residues present in radius of 15 Å from the metal ion center. In this way, the interactions between α and β subunits of the homodimer were included. Inside the considered sphere, an outer shell of residues with a thickness of 2 Å was fixed, and only the inner 13 Å shell was allowed to move during the QM/MM geometry optimizations. This strategy is commonly used to avoid drifting through multiple minima unrelated to the reaction coordinate. This model includes in the MM region also a number of water molecules (20) present in the crystallographic structure. The obtained model consists of 2,154 atoms with 118 atoms in QM region (**Figure 1**).

ONIOM-2

A rectangular box was used to solvate the system up to 12.0 Å of the metal center. During the optimizations, all the water molecules and protein atoms in the 18 Å from the active site were kept frozen, as proposed by a recent work (Medina et al., 2017). The final model contains 11,895 atoms with 118 atoms of QM region. In this case, the MM region includes the whole protein and a number of water molecules within 5 Å around of catalytic domain as depicted in **Figure 1**.

RESULTS AND DISCUSSION

The reaction can follow two paths with the formation of $\mathbf{CO_2}$ or $\mathbf{HCO_3^-}$ products (see **Figure 2**). After the formation of the **ES**, the reaction proceeds with the proton transfer from Asp314 to C5 of the substrate generating the **INT1** species, that acts as common intermediate for the formation of **EP_I** or **EP_II** complexes in which $\mathbf{CO_2}$ or $\mathbf{HCO_3^-}$ product should be released. In both decarboxylation pathways, it is clear that the enzyme must generate an adjacent electron sink (such as the ketone carbonyl C4 since the formation of the new carbon–hydrogen bond) to stabilize the incipient carbanion at C5 prior to decarboxylation. This mechanism corresponds to that explored in the recent combined experimental and theoretical work (Sheng et al., 2017) where the membrane inlet mass spectrometry (MIMS) based assay is applied to study the LigW mechanism. The above-mentioned MIMS-based strategy (Sheng et al., 2017) was able to establish $\mathbf{CO_2}$ and not $\mathbf{HCO_3^-}$ as reaction product. We have considered also the path for the bicarbonate release but our calculated PESs with the three models used give very high energetic barriers (see Table S1) that are not compatible with the enzymatic kinetics.

All the obtained PESs with the used models are depicted in **Figure 3**. Those concerning the QM one will be compared with

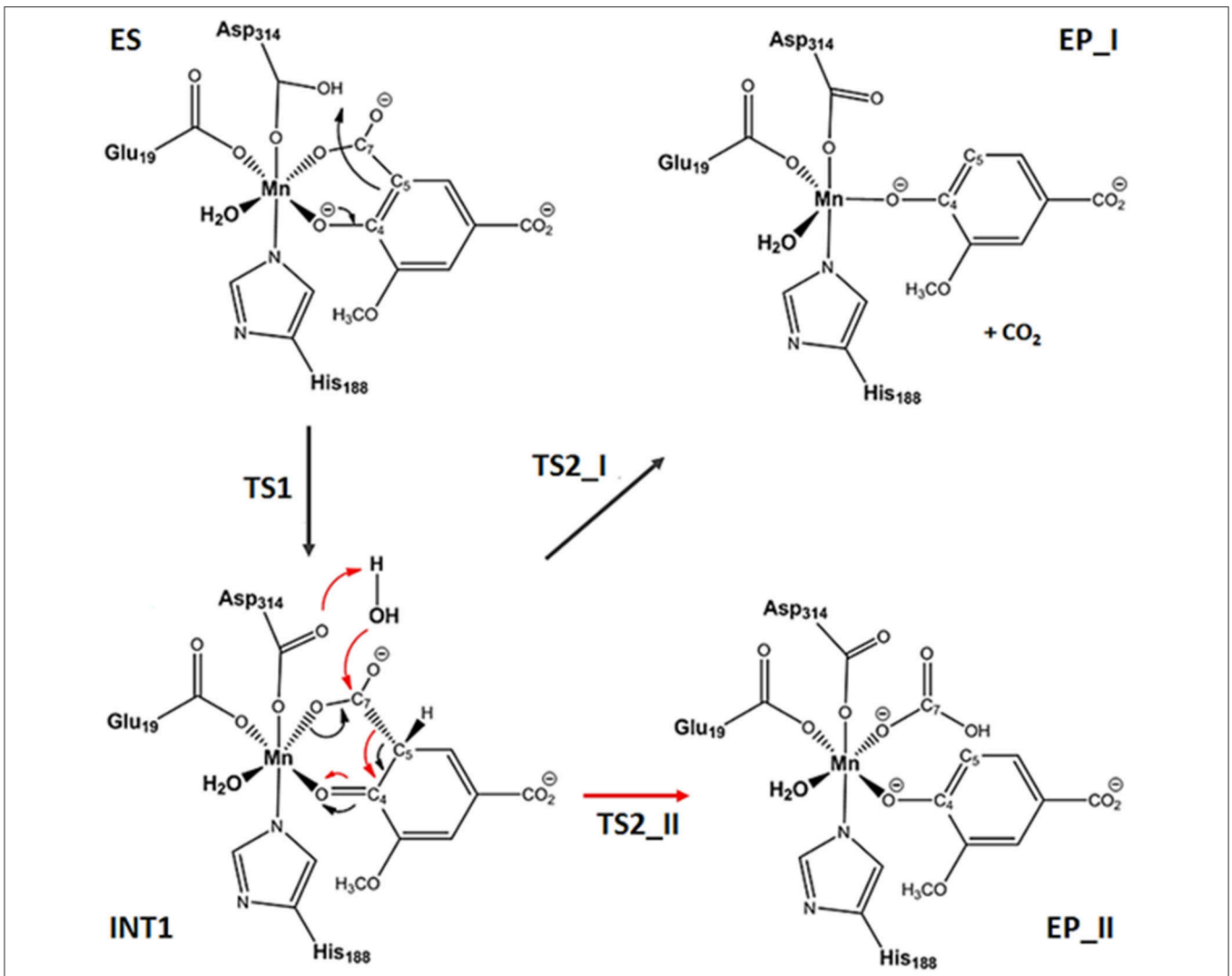


FIGURE 2 | Reaction mechanism of the catalyzed by LigW. For the decarboxylation step, CO_2 and HCO_3^- formations occur through TS2_I and TS2_II transition states.

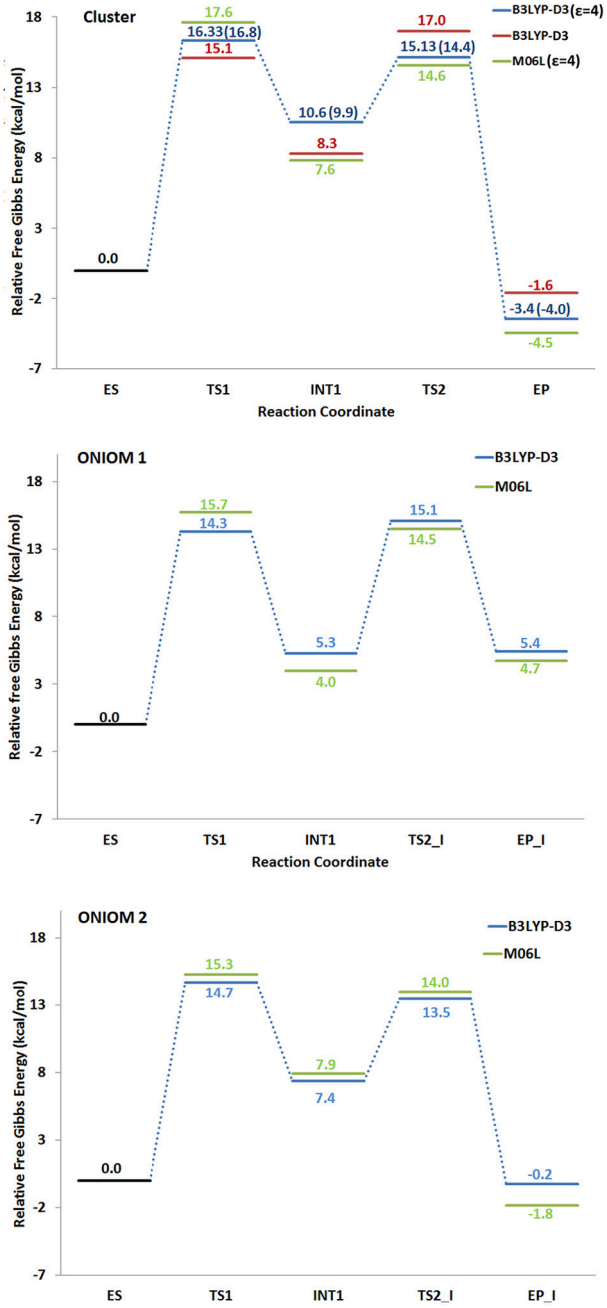


FIGURE 3 | Potential energy surfaces obtained with different density functionals. The numbers in paranthesis correspond to those obtained in a previous larger QM cluster study (Sheng et al., 2017).

the values arising from the previous larger QM-cluster model study (Sheng et al., 2017). B3LYP optimized structures obtained employing the ONIOM-2 model of all the species are given in **Figure 4** while that for the QM and ONIOM-1 models are given in Figures S2, S3.

In **ES** complex, the Asp314, as in the original X-ray structure, is oriented in a suitable way to deliver the proton to C5 of the substrate (H_{Asp314}-C5 3.103 Å). **w2** and **w3** water molecules originally bonded to the metal ion and displaced upon the substrate entrance, lie in proximity to the reaction site establishing H- bonds network with the surrounding amino acid residues (see Figure S4). The bond lengths in the active site of the present (126 atoms) and previous (Sheng et al., 2017) larger (308 atoms) QM cluster study agree very well.

The formation of **INT1** takes place through the transition state **TS1** that describes the proton transfer from the Asp314 to the carbon atom of 5-CV. The related imaginary frequency ($669i$ cm^{-1}) well accounts for this process since it is associated to the stretching vibrational motions of the proton transfer (O–H and H–C5). The analysis of the **TS1** optimized structure (**Figure 4** for ONIOM-2 and Figure S2 for QM cluster) reveals that the formation of the C5-H bond (1.303 Å) is more advanced in the case of ONIOM-2 calculation. In fact, the breaking bond between hydrogen and oxygen of Asp314 (1.611 Å) is more elongated than the usual sp^3 O-H bond. Furthermore, a major distortion of the $-COO^-$ moiety out of plane of the phenyl ring of the substrate can be observed (76 degrees in ONIOM-2 vs. 19 degrees in QM cluster). These geometrical differences may be responsible from the slight variations in the **TS1** barrier (14.7 kcal/mol and 16.3 kcal/mol for ONIOM-2 and QM cluster, respectively). **INT1** (**Figure 4**) is characterized by a C5-C7 single bond with a distance slightly elongated (1.613 Å) with respect to the single canonical bond (C-C) and a sp^3 C5 hybridized prone for the subsequent decarboxylation step. The barrier for the $\mathbf{CO_2}$ formation (**TS2_I**) is calculated to be 13.4 kcal/mol above **ES** complex, (only 6 kcal/mol relative to the **INT1**). The present QM cluster model obtains this barrier as 15.1 kcal/mol, analogous to the result (14.4 kcal/mol) of the previous (Sheng et al., 2017) cluster study with larger QM size.

The **TS2_I** is characterized by the C5-C7 distance of 1.853 Å associated with a relative imaginary C-C stretching frequency of $129i$ cm^{-1} (**Figure 4**). The already formed carbon dioxide is still coordinated to the metal ion (2.240 Å) and the manganese ion is still hexa-coordinate in octahedral geometry fashion (**Figure 4**). This topology is present in all our used models and in the previous larger QM cluster. (Sheng et al., 2017) At the end of the decarboxylation process, one molecule of carbon dioxide is released and the **EP_I** complex is generated (see **Figure 4**). The manganese ion assumes a trigonal bipyramidal geometry due to the loss of the sixth ligand ($\mathbf{CO_2}$). The created vacancy will be filled by one of the two water molecules present in active site (**w2** and **w3**) and essential to restore the catalytic cycle. ONIOM-2 offers a better value of the reaction energy (0.2 kcal/mol below the **ES** complex, see **Figure 3**) while at QM level it is exergonic (−3.5 kcal/mol, see **Figure 3**). In order to verify the role of the bulk potential on the cluster model, single point computations were performed on the previous optimized structures removing all the environmental effects. Results, reported in **Figure 3**, show that the PES behavior is almost retained. The largest effect (−2.3 kcal/mol) concerns the INT1 species.

NBO charges trend illustrated in **Figure 5** confirms the nonoxidative nature of the decarboxylation process as evinced

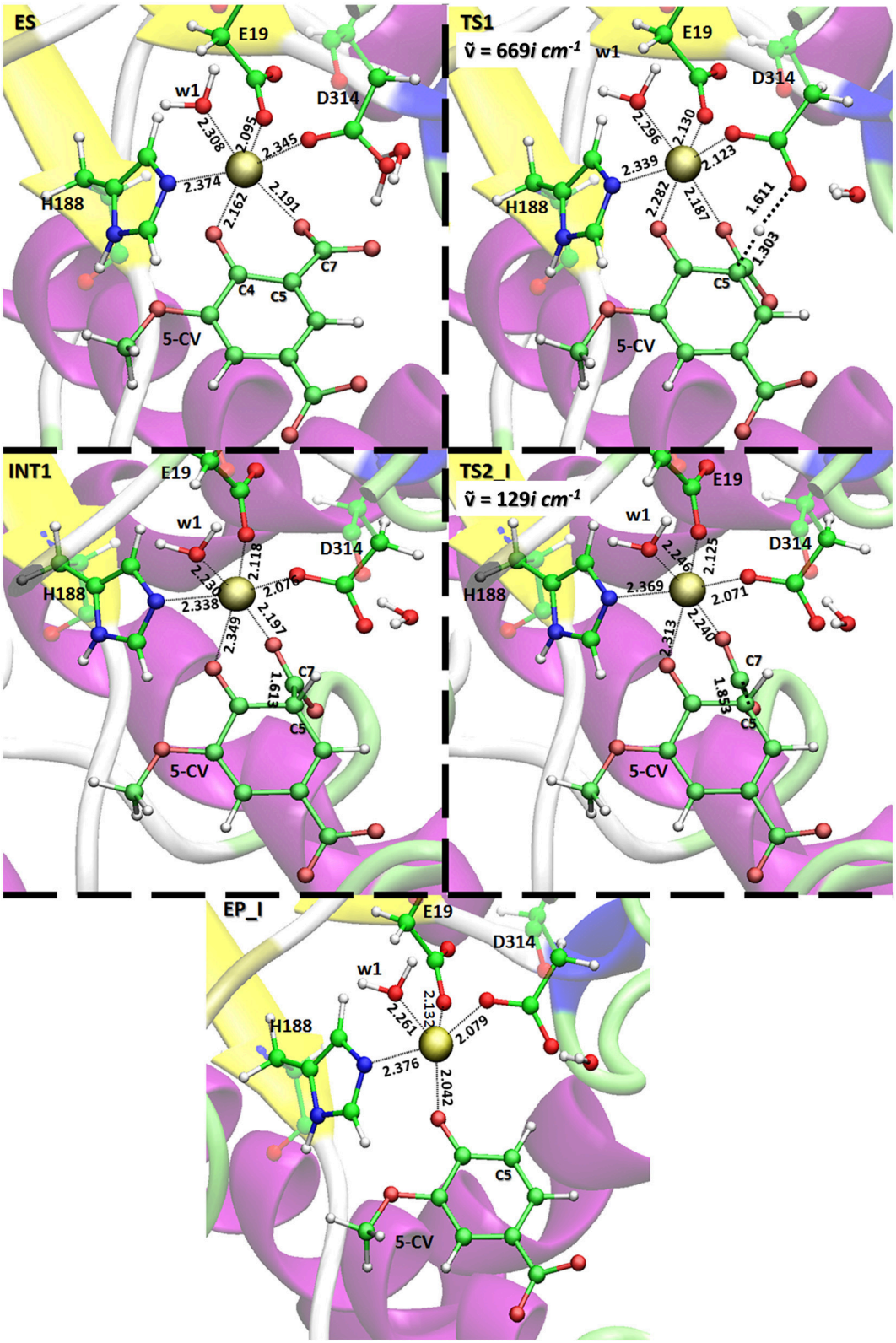


FIGURE 4 | Optimized QM layer in ball and sticks (ONIOM 2 model) of all species present on the potential energy surface obtained with the B3LYP/6-31G(d,p)|SDD:FF99SB level. For clarity, only the amino acid residues of the inner coordination shell of the metal center are retained. The distances are indicated in Å. Imaginary frequencies of the transition states are also reported.

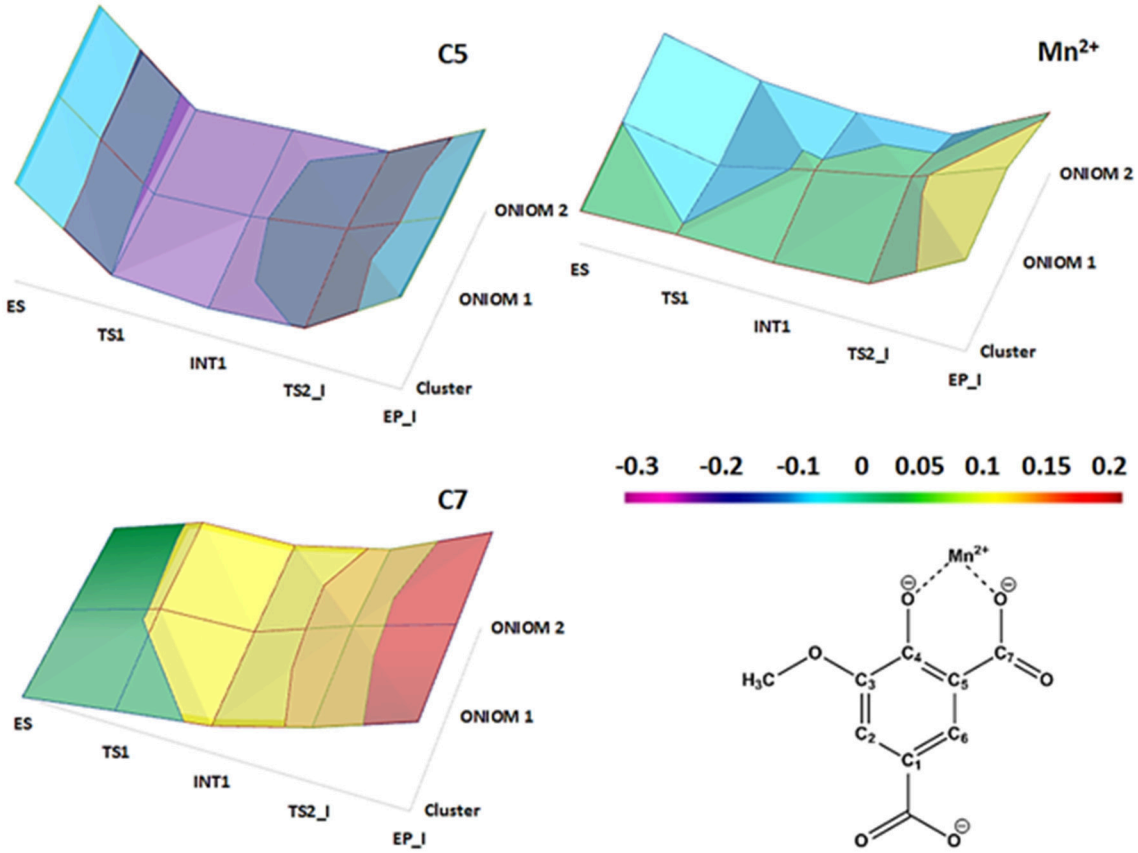


FIGURE 5 | NBO charge distributions on the C5, C7, and Mn^{2+} centers directly involved in the reaction steps for the three models used.

from the average value of the charges of the Mn^{2+} (1.117 *e*), the C5 (−0.417 *e*) and the C7 (0.918 *e*) atoms in all the species intercepted on the PES. From the Figure S5, it can be also evidenced that the nonbonded interactions (characterizing the amino acid residues of the inner coordination shell with the metal ion) as well as the stacking interactions between the substrate (product) and Tyr317 are retained during the reaction.

All the models propose the **TS1** which describes the formation of the C5-protonated intermediate, as the rate limiting step (14.7 kcal/mol). Based on the experimental k_{cat} value of 27 s^{-1} for *Sphingomonas paucimobilis* LigW (Sheng et al., 2017), the reaction barrier is expected to be ∼16 kcal/mol. The closeness of the experimental estimate of the reaction barrier and computational **TS1** barrier suggest the appropriateness of the present and previous computational protocols.

The optimized species intercepted along the PES for the bicarbonate release (step II) are shown in Figure S6. The **w3** molecule comes into play in the reaction since it performs a nucleophilic attack on the carbon (C7) (O_{w3} – C7 distance of 1.944 Å) for generating the $\mathbf{HCO_3^-}$ species and simultaneously donating a proton to Asp314 (H_w – O_{Asp314} distance of 1.532 Å). The obtained energy barrier is 30.3 kcal/mol (see Table S1).

CONCLUSION

In this work, we have investigated the reaction mechanism of LigW by using three different models and two exchange correlation functionals. This allowed us to assess the influence of the employed model on the computed structures and energetics compared with available experimental data. The models include full structure, its partial solvation, and reactive center with the rest represented by a bulk potential including geometrical restraints at the border. Since the results of these three models, previous larger QM cluster and experimental studies are consistent to each other, the amino acids and waters outside the reactive center act on the reaction energetics in an average way for the present enzyme system. A similar behavior was also observed in many other enzymes (Himo, 2006, 2017; Blomberg et al., 2014). However, one should keep in mind that every enzyme system acts differently and thus one should avoid the generalization of the result despite its validity on a large variety of enzyme systems.

AUTHOR CONTRIBUTIONS

MP, TM, and NR have analyzed the results, edit and reviewed equally the manuscript. MP, TM, and NR approved it for publications.

ACKNOWLEDGMENTS

Financial support from the Università degli Studi della Calabria -Dipartimento di Chimica e Tecnologie Chimiche (CTC) is acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2018.00249/full#supplementary-material

REFERENCES

Alberto, M. E., Marino, T., Ramos, M. J., and Russo, N. (2010). Atomistic details of the Catalytic Mechanism of Fe(III)-Zn(II) Purple Acid Phosphatase. *J. Chem. Theory Comput*. 6, 2424–2433. doi: 10.1021/ct100187c

Amata, O., Marino, T., Russo, N., and Toscano, M. (2011a). Catalytic activity of a ζ-class zinc and cadmium containing carbonic anhydrase. Compared work mechanisms. *Phys. Chem. Chem. Phys*. 13, 3468–3477. doi: 10.1039/c0cp01053g

Amata, O., Marino, T., Russo, N., and Toscano, M. (2011b). A proposal for mitochondrial processing peptidase catalytic mechanism. *J. Am. Chem. Soc.* 133, 17824–17831. doi: 10.1021/ja207065v

Andrae, D., Haußermann, H., Dolg, M., Stoll, H., and Preuß, H. (1990). Energy-adjusted *ab initio* pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta* 77, 123–141, doi: 10.1007/BF01114537

Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem*. 97, 10269–10280, doi: 10.1021/j100142a004

Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 98, 5648–5652. doi: 10.1063/1.464913

Blomberg, M. R. A., Borowski, T., Himo, F., Liao, R. Z., and Siegbahn, P. E. M. (2014). Quantum chemical studies of mechanisms for metalloenzymes. *Chem. Rev*. 114, 3601–3658. doi: 10.1021/cr400388t

Duarte, M., Ferreira-da-Silva, F., Lünsdorf, H., Junca, H., Gales, L., Pieper, D. H., et al. (2011). Gulosibacter molinativorax ON4T molinate hydrolase, a novel cobalt-dependent amidohydrolase. *J. Bacteriol*. 193, 5810–5816. doi: 10.1128/JB.05054-11

Elias, M., Dupuy, J., Merone, L., Mandrich, L., Porzio, E., Moniot, S., et al. (2008). Structural basis for natural lactonase and promiscuous phosphotriesterase activities. *J. Mol. Biol*. 379, 1017–1028. doi: 10.1016/j.jmb.2008.04.022

Faponle, A. S., Quesne, M. G., and de Visser, S. P. (2016). Origin of the regioselective fatty-acid hydroxylation versus decarboxylation by a cytochrome P450 peroxygenase: what drives the reaction to biofuel production? *Chem. Eur. J*. 22, 5478–5483. doi: 10.1002/chem.201600739

Gerlt, J. A., and Babbitt, P. C. (2001). Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annu. Rev. Biochem*. 70, 209–246. doi: 10.1146/annurev.biochem.70.1.209

Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem*. 32, 1456–1465. doi: 10.1002/jcc.21759

Himo, F. (2006). Quantum chemical modeling of enzyme active sites and reaction mechanisms. *Theor. Chem. Acc*. 116, 232–240. doi: 10.1007/s00214-005-0012-1

Himo, F. (2017). Recent trends in quantum chemical modeling of enzymatic reactions. *J. Am. Chem. Soc*. 139, 6780–6786. doi: 10.1021/jacs.7b02671

Jackson, C. J., Liu, J. W., Coote, M. L., and Ollis, D. L. (2005). The effects of substrate orientation on the mechanism of a phosphotriesterase. *Org. Biomol. Chem*. 3, 4343–4350. doi: 10.1039/b512399b

Khurana, J. L., Jackson, C. J., Scott, C., Pandey, G., Horne, I., Russell, R. J., et al. (2009). Characterization of the phenylurea hydrolases A and B: founding members of a novel amidohydrolase subgroup. *Biochem. J*. 418, 431–441. doi: 10.1042/BJ20081488

Lee, C., Yang, W., and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37, 785–789. doi: 10.1103/PhysRevB.37.785

Liao, R. Z., Yu, J. G., and Himo, F. (2010). Mechanism of tungsten-dependent acetylene hydratase from quantum chemical calculations. *Proc. Natl. Acad. Sci. U.S.A*. 107, 22523–22527. doi: 10.1073/pnas.1014060108

Liu, A., and Zhang, H. (2006). Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry* 45, 10407–10411. doi: 10.1021/bi061031v

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput*. 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255

Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* 113, 6378–6396. doi: 10.1021/jp810292n

Medina, F. E., Neves, R. P. P., Ramos, M. J., and Fernandes, P. A. (2017). A QM/MM study of the reaction mechanism of human β-ketoacyl reductase. *Phys. Chem. Chem. Phys*. 19, 347–355. doi: 10.1039/C6CP07014K

Merz, K. M. (2014). Using quantum mechanical approaches to study biological systems. *Acc. Chem. Res*. 47, 2804–2811. doi: 10.1021/ar5001023

Piazzetta, P., Marino, T., Russo, N., and Salahub, D. R. (2017). The role of metal substitution in the promiscuity of natural and artificial carbonic anhydrases. *Coord. Chem. Rev*. 345, 73–85. doi: 10.1016/j.ccr.2016.12.014

Prejanò, M., Marino, T., Rizzuto, C., Madrid, J. C. M., Russo, N., and Toscano, M. (2017a). Reaction mechanism of low-spin iron(III)- and cobalt(III)-containing nitrile hydratases: a quantum mechanics investigation. *Inorg. Chem*. 56, 13390–13400. doi: 10.1021/acs.inorgchem.7b02121

Prejanò, M., Marino, T., and Russo, N. (2017b). How can methanol dehydrogenase from *Methylacidiphilum fumariolicum* work with the alien Ce^{III} Ion in the active center? A theoretical study. *Chem. Eur. J*. 23, 8652. doi: 10.1002/chem.201700381

Quesne, M. G., Borowski, T., and de Visser, S. P. (2016). Quantum mechanics/molecular mechanics modeling of enzymatic processes: caveats and breakthroughs. *Chem. Eur. J*. 22, 2562–2581. doi: 10.1002/chem.201503802

Ramos, M. J., and Fernandes, P. A. (2008). Computational enzymatic catalysis. *Acc. Chem. Res*. 41, 689–698, doi: 10.1021/ar7001045

Ryde, U. (2016). QM/MM calculations on proteins. *Methods Enzymol*. 577, 119–158. doi: 10.1016/bs.mie.2016.05.014

Ryde, U. (2017). How many conformations need to be sampled to obtain converged QM/MM energies? The Curse of Exponential Averaging. *J. Chem. Theory Comput*. 13, 5745–5752. doi: 10.1021/acs.jctc.7b00826

Seibert, C. M., and Raushel, F. M. (2005). Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* 44, 6383–6391. doi: 10.1021/bi047326v

Senn, H. M., and Thiel, W. (2009). QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed*. 48, 1198–1229. doi: 10.1002/anie.200802019

Shapir, N., Pedersen, C., Gil, O., Strong, L., Seffernick, J., Sadowsky, M. J., et al. (2006). TrzN from Arthrobacter aurescens TC1 Is a zinc amidohydrolase. *J. Bacteriol*. 188, 5859–5864. doi: 10.1128/JB.00517-06

Sheng, X., Zhu, W., Huddleston, J., Xiang, D. F., Raushel, F. M., Richards, N. G. J., et al. (2017). A combined experimental-theoretical study of the LigW-catalyzed decarboxylation of 5-carboxyvanillate in the metabolic pathway for lignin degradation. *ACS Catal*. 7, 4968–4974. doi: 10.1021/acscatal.7b01166

Siegbahn, P. E. M., and Himo, F. (2011). The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci*. 1, 323–336. doi: 10.1002/wcms.13

Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996). ONIOM: a multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. a test for diels–alder reactions and $Pt(P(t\text{-}Bu)_3)_2 + H_2$ oxidative addition. *J. Phys. Chem*. 100, 19357–19363. doi: 10.1021/jp962071j

Tobimatsu, Y., Chen, F., Nakashima, J., Escamilla-Treviño, L. L., Jackson, L., Dixon, R. A., and et al. (2013). Coexistence but independent biosynthesis of catechyl and guaiacyl/syringyl lignin polymers in seed coats. *Plant. Cell*. 25, 2587–2600. doi: 10.1105/tpc.113.113142

Vladimirova, A., Patskovsky, Y., Fedorov, A. A., Bonanno, J. B., Fedorov, E. V., Toro, R., et al. (2016). Substrate distortion and the catalytic reaction

mechanism of 5-carboxyvanillate decarboxylase. *J. Am. Chem. Soc*. 138, 826–836. doi: 10.1021/jacs.5b08251

Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, J. A. Jr., Morokuma, K., et al. (2006). Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput*. 2, 815–826. doi: 10.1021/ct050289g

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *J. Comp. Chem*. 25, 1157–1174. doi: 10.1002/jcc.20035

Warshel, A. (1991). *In Computer Modeling of Chemical Reactions in Enzymes and Solutions*. New York, NY: Wiley.

Warshel, A., and Levitt, M. (1976). Theoretical studies of enzymatic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol*. 103, 227–249. doi: 10.1016/0022-2836(76)90311-9

Zhao, Y., and Truhlar, D. G. (2006). A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys*. 125, 1–18. doi: 10.1063/1.2370993

Zhao, Y., and Truhlar, D. G. (2008). The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc*. 120, 215–241. doi: 10.1007/s00214-007-0310-x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Prejanò, Marino and Russo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Supplementary Material

QM cluster or QM/MM in computational enzymology: the test case of LigW-decarboxylase

Mario Prejanò[1],Tiziana Marino[1*], Nino Russo[1]

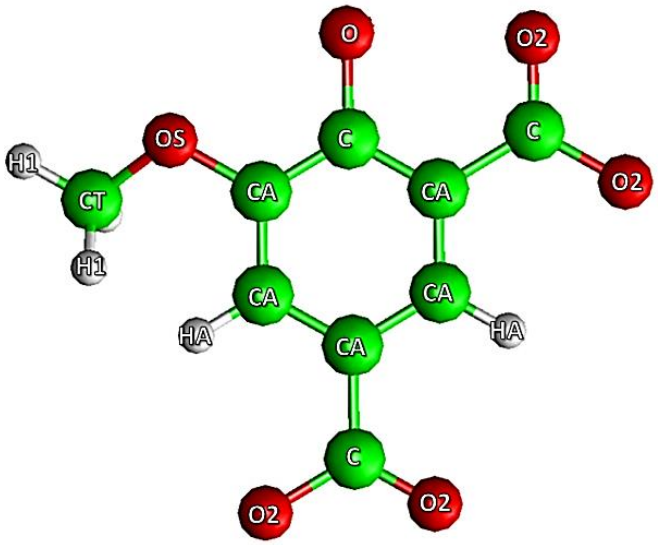
*** Correspondence:** Tiziana Marino: tiziana.marino65@unical.it

Supplementary Data

Tab. S1 Energy values of all the species present on the potential energy surfaces of two explored paths.

	Cluster		ONIOM 1		ONIOM 2	
	SMD(ε=4)6-311+G(2d,2p)|SDD		6-311+G(2d,2p)|SDD: amber		6-311+G(2d,2p)|SDD:amber	
	B3LYP-D3	M06-L	B3LYP-D3	M06-L	B3LYP-D3	M06-L
ES	0.0	0.0	0.0	0.0	0.0	0.0
TS1	16.3	17.6	14.3	15.7	14.7	15.3
INT1	10.6	7.9	5.3	4.0	7.4	7.9
TS2_I	15.1	14.6	15.1	14.5	13.5	14.0
EP_I	-3.4	-4.5	5.4	4.7	-0.2	-1.8
TS2_II	34.2	41.6	43.9	46.3	30.3	31.9
EP_II	3.9	-7.9	12.6	11.7	-6.9	-4.8

Tab. S2. Calculated parameters for 5-carboxyvanillate.



Atomtype	Charge
CA	-0.266
OS	-0.443
CT	0.135
H1	0.007
HA	0.120
C	0.974
O2	-0.896
O	-0.786

Bond	K_i / kcal mol^{-1} Å^{-2}	l_0/Å
CA-CA	478.40	1.387
CA-C	449.90	1.406
CA-HA	344.30	1.087
CA-OS	392.60	1.357
OS-CT	301.50	1.439
CT-H1	337.30	1.092
C-O2	648.00	1.214
C-O	648.00	1.214

Angle	K_{an}/ Kcal mol^{-1} rad^{-2}	θ_0/deg
CA-CA-CA	67.180	119.970
CA-CA-HA	50.300	119.700
CA-C-O2	72.770	119.120
CA-CA-C	67.930	120.700
CA-CA-OS	71.040	121.890
CA-OS-CT	64.210	112.090
CA-C-CA	67.170	116.780
CA-C-O	72.770	119.120
OS-CA-C	70.229	117.340
OS-CT-H1	50.870	108.700
H1-CT-H1	39.430	108.350
O2-C-O2	78.170	130.380
C-CA-C	66.570	118.880

Dihedral	d	V_n/	γ/deg	n

		Kcal mol^{-1} rad^{-2}		
CA-CA-CA-OS	1	6.650	180.000	2.000
CA-CA-CA-C	1	6.650	180.000	2.000
CA-CA-C-O2	1	2.175	180.000	-2.000
CA-CA-C-O2	1	0.300	0.000	3.000
CA-CA-CA-HA	1	6.650	180.000	2.000
CA-CA-CA-CA	1	3.625	180.000	2.000
CA-CA-OS-CT	1	1.050	180.000	2.000
CA-CA-C-CA	1	2.175	180.000	2.000
CA-CA-C-O	1	2.175	180.000	-2.000
CA-CA-C-O	1	0.300	0.000	3.000
CA-OS-CT-H1	1	0.383	0.000	3.000
CA-C-CA-C	1	2.175	180.000	2.000
OS-CA-CA-HA	1	6.650	180.000	2.000
OS-CA-C-CA	1	2.175	180.000	2.000
OS-CA-C-O	1	2.175	180.000	2.000
CT-OS-CA-C	1	1.050	180.000	2.000
HA-CA-CA-C	1	6.650	180.000	2.000
C-CA-C-O	1	2.175	180.000	2.000
O2-C-CA-C	1	2.175	180.000	2.000

Improper	V_n/ Kcal mol^{-1} rad^{-2}	γ/deg	n
C-CA-CA-CA	1.1	180.0	2.0
CA-CA-CA-HA	1.1	180.0	2.0
C-CA-CA-OS	1.1	180.0	2.0
CA-O2-C-O2	1.1	180.0	2.0
C-C-CA-CA	1.1	180.0	2.0
CA-CA-C-O	1.1	180.0	2.0

VdW	r_0 /Å	ε/ kcal mol^{-1}
CA	1.9080	0.0860
OS	1.6837	0.1700
CT	1.9080	0.1094
H1	1.4870	0.0157
HA	1.4870	0.0157
C	1.9080	0.0860
O2	1.6612	0.2100
O	1.6612	0.2100

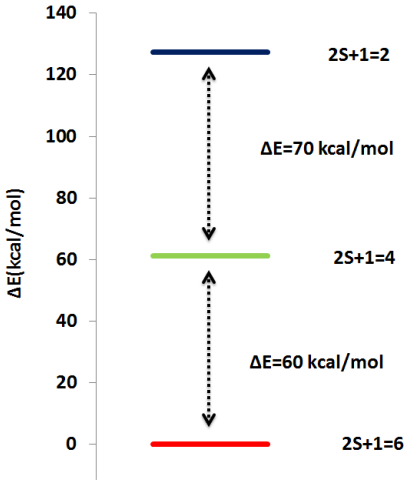


Fig. S1 Energetic behavior of the ES species at the considered spin multiplicities.

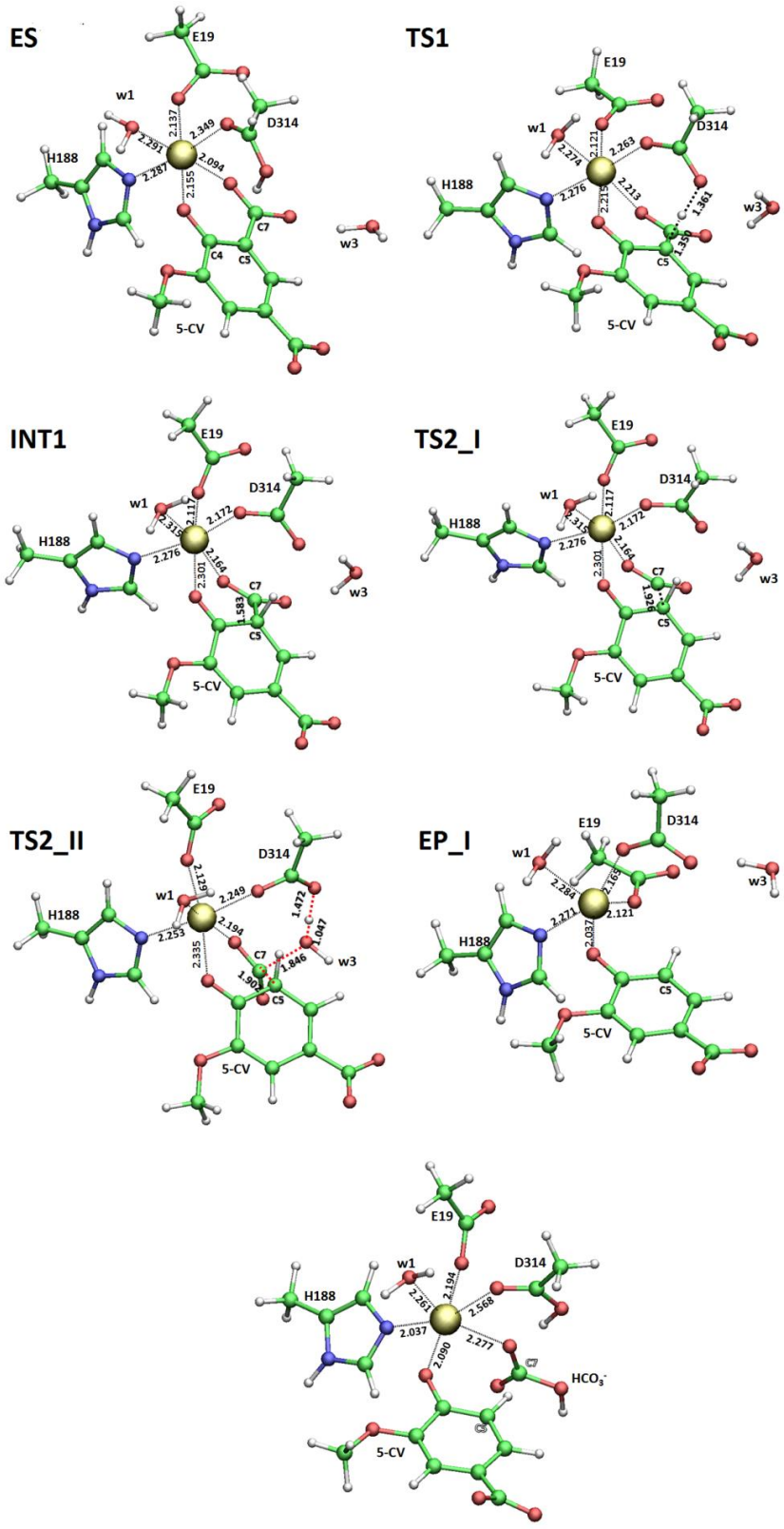


Fig. S2 B3LYP/6-31+G(d,p) (SDD for Mn) optimized geometries in QM-cluster for both explored paths.

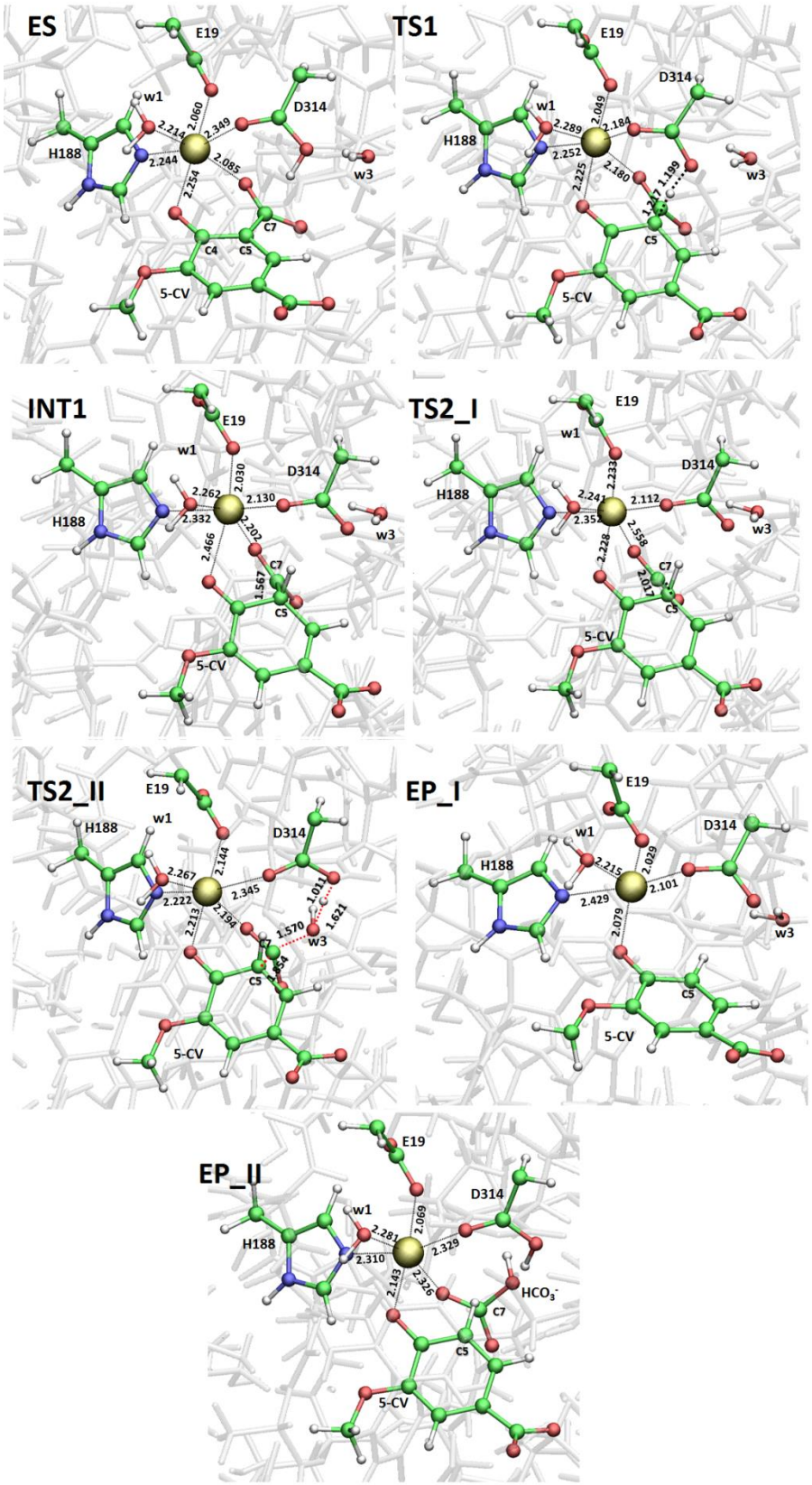


Fig. S3 ONIOM 1 optimized geometries for both explored paths.

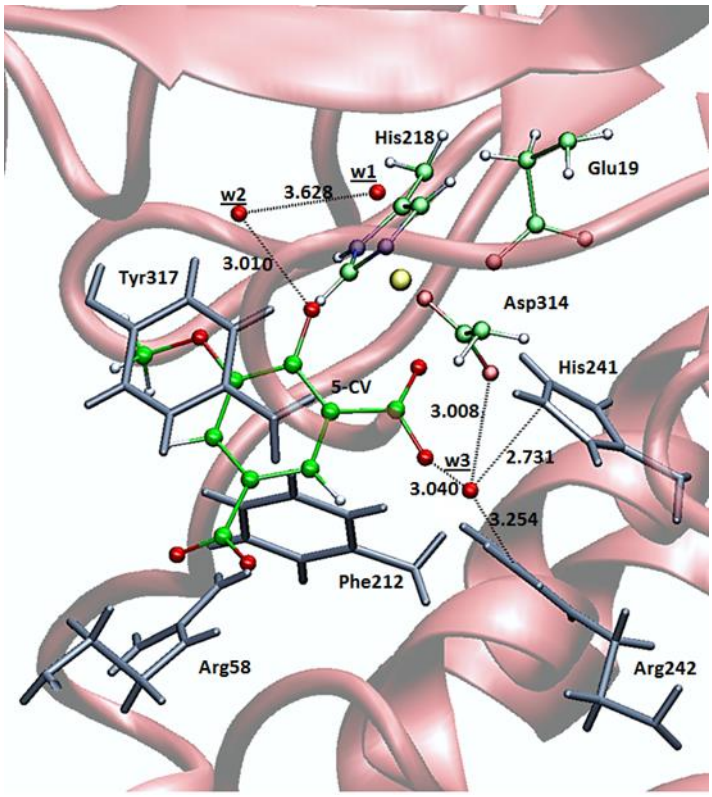


Fig. S4 Water molecules (**w2** and **w3**) implicated in the hydrogen bonds network in ES complex (ONIOM-2).

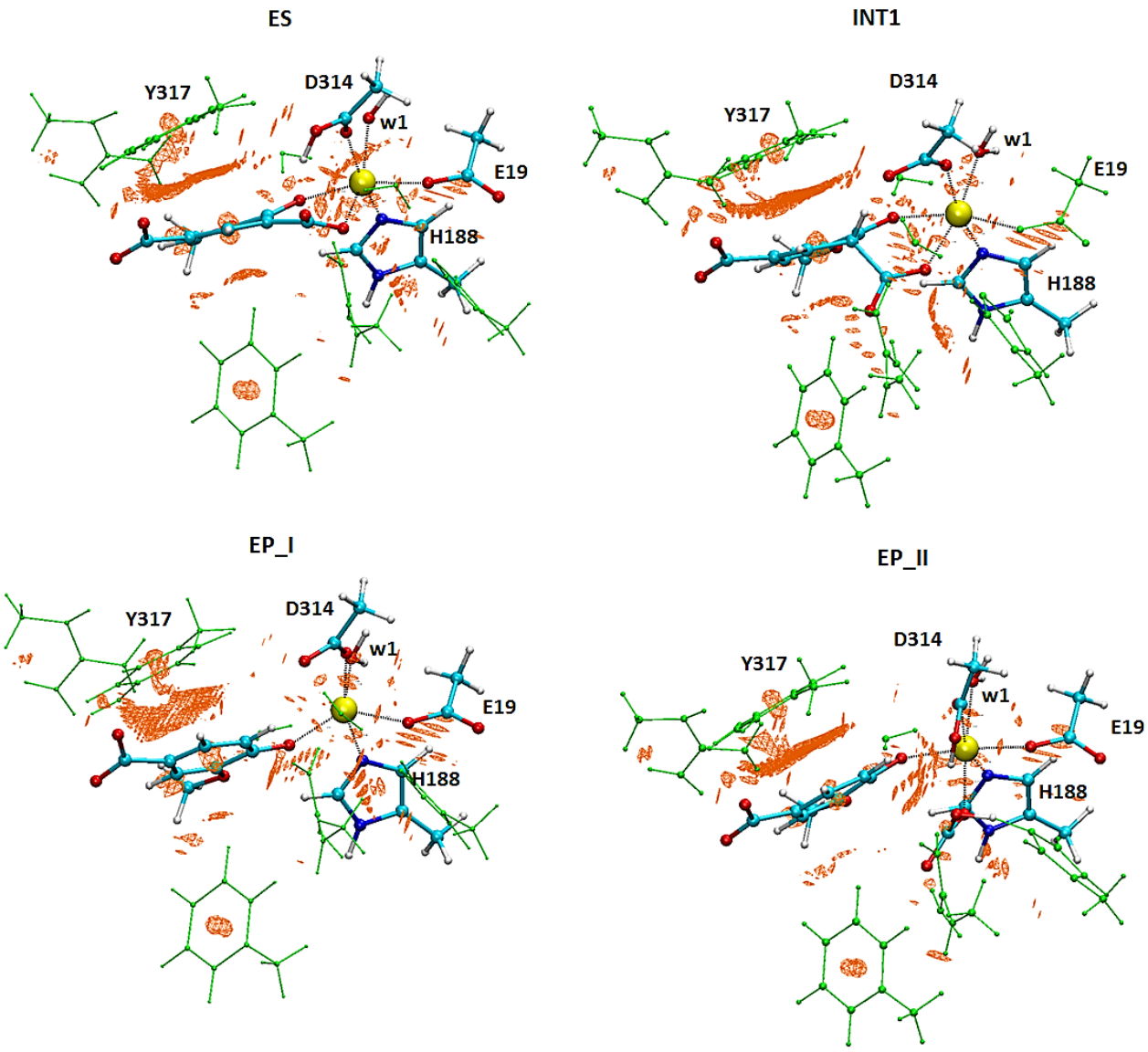


Figure S5. Nonbonding interactions plot (orange isosurfaces) calculated for the minima intercepted on the PES.

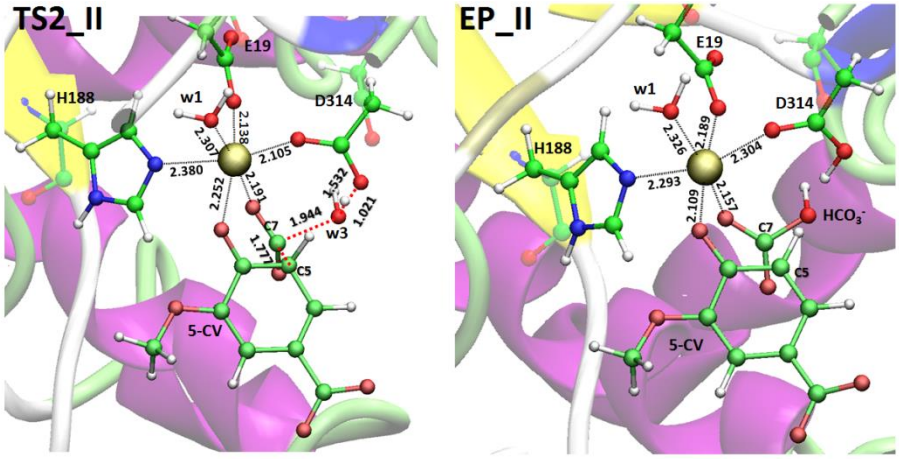


Fig. S6. ONIOM 2 optimized geometries of the characterized species present on the HCO_3^- formation path.

Paper V

"On the inhibition mechanism of glutathione transferase P1 by piperlongumine. Insight from theory"

Mario Prejanò, Tiziana Marino, Nino Russo

Manuscript accepted

On the inhibition mechanism of glutathione transferase P1 by piperlongumine. Insight from theory

Tiziana Marino[1, 2*], Mario Prejanò[1, 2], Nino Russo[1, 2]

[1]Università della Calabria, Italy, [2]Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Italy

Submitted to Journal:
Frontiers in Chemistry

Specialty Section:
Theoretical and Computational Chemistry

Article type:
Original Research Article

Manuscript ID:
423376

Received on:
05 Sep 2018

Revised on:
23 Nov 2018

Frontiers website link:
www.frontiersin.org

In review

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

MP, TM, and NR have analyzed the results, edit and reviewed equally the manuscript. MP, TM, and NR approved it for publications.

Keywords

Glutathione S-transferase, Piperlongumine, Hydrolysis mechanism, Inhibition mechanism, MD, DFT, QM, QMMM

Abstract

Word count: 136

Piperlongumine (PL) is an anticancer compound whose activity is related to the inhibition of human glutathione transferase of pi class (GSTP1) overexpressed in cancerous tumors and implicated in the metabolism of electrophilic compounds. In the present work, the inhibition mechanism of hydrolyzed piperlongumine (hPL) has been investigated employing QM and QM/MM levels of theory. The potential energy surfaces (PESs) underline the contributions of Tyr residue close to G site in the catalytic pocket of the enzyme. The proposed mechanism occurs through a one-step process represented by the nucleophilic addition of the glutathione thiol to electrophilic species giving rise to the simultaneous C-S and H-C bonds formation. Both the used methods give barrier heights (19.8 and 21.5 kcal mol-1 at QM/MM and QM, respectively) close to that experimentally measured for the C-S bond formations. (23.8 kcal mol-1)

In review

Ethics statements

(Authors are required to state the ethical considerations of their study in the manuscript, including for cases where the study was exempt from ethical approval procedures)

Does the study presented in the manuscript involve human or animal subjects: No

On the inhibition mechanism of glutathione transferase P1 by piperlongumine. Insight from theory

Mario Prejanò, Tiziana Marino*, Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via P. Bucci, I-87036-Arcavacata di Rende, Italy

***Correspondence:**
Tiziana Marino

tiziana.marino65@unical.it

glutathione S-transferase, piperlongumine, hydrolysis mechanism, inhibition mechanism, MD DFT, QM, QMMM

In review

Abstract

Piperlongumine (PL) is an anticancer compound whose activity is related to the inhibition of human glutathione transferase of pi class (GSTP1) overexpressed in cancerous tumors and implicated in the metabolism of electrophilic compounds. In the present work, the inhibition mechanism of hydrolyzed piperlongumine (hPL) has been investigated employing QM and QM/MM levels of theory. The potential energy surfaces (PESs) underline the contributions of Tyr residue close to *G site* in the catalytic pocket of the enzyme. The proposed mechanism occurs through a one-step process represented by the nucleophilic addition of the glutathione thiol to electrophilic species giving rise to the simultaneous C-S and H-C bonds formation. Both the used methods give barrier heights (19.8 and 21.5 kcal mol^{-1} at QM/MM and QM, respectively) close to that experimentally measured for the C-S bond formations. (23.8 kcal mol^{-1})

1 Introduction

Glutathione-S-transferases (GSTs) is a ubiquitous family of multifunctional enzymes of phase II detoxification system that conjugate reactive substrates with reduced tripeptide glutathione (GSH) in most cells, especially those in the liver and kidney. (Hayes et al., 2005; Oakley et al., 2005; Stoddard et al., 2017) In particular, they catalyze the nucleophilic attack of the thiol group arising from cysteine residue (Cys) of the GSH on electrophilic substrates leading to formation of conjugates, that are less toxic and more water-soluble than the parent species, facilitating their elimination from cells. (Wang et al, 2017; Townsend and Tew, 2003; Broxterman et al, 1995) Their role in protecting the cells from oxidative attack, in association with their overexpression in many cancer cells, makes them good candidates as cancer biomarkers. (McIlwain et al, 2006; Lo and Ali-Osman, 2007) Furthermore, glutathione-S-transferases are associated with multidrug resistance of tumor cells and are involved in drug detoxification and in apoptosis control. (Townsend and Tew., 2003; Mejerman et al., 2008) Mammalian cytosolic GSTs isoenzymes belong to different families or classes (alpha, mu, pi, theta, kappa, sigma, zeta, and omega) (Armstrong, 1997; Sheehan et al., 2001; Wilce and Parker, 1994) based on their molecular masses, isoelectric points and other properties. Every isoenzyme subunit contains an active site entailing a binding site for the cofactor GSH (*G-site*) and one for the electrophilic substrate (*H-site*). (Wilce and Parker, 1994; Dirr et al., 1994) In particular,

the Glutathione S-transferase P1 (GSTP1) is overexpressed in different human malignancies affecting important organs as lung, colon, stomach, kidney, ovary, mouth and testis. (Inoue et al., 1995; Ruiz-Gomez et al., 2000; Okuyama et al., 1994; Grignon et al, 1994; Green et al., 1993; Zhang et al., 1994; Katagiri et al, 1993) This overexpression has been linked to acquire multidrug resistance to chemotherapeutic agents (cisplatin, chlorambucil, and ethacrynic acid). (Mejerman et al, 2008; Ban et al., 1996; Oakley et al., 1997; Karpusas et al., 2013; Pei et al., 2013; Perperopoulou et al, 2018) GSTP1 has additional role in maintaining the cellular redox state (Tew, 2007) and ‘‘nonenzymatic’’ antiapoptotic activity through its interaction with the *c-Jun* NH_2-terminal kinase (JNK), a key enzyme implicated into the apoptotic cascade. (Adler et al., 1999; Wang et al.; 2001) For these reasons, GSTP1 is considered as a promising target for inactivation in cancer treatment and numerous researchers have spent considerable effort to propose potent inhibitors of this enzyme. (Raj et al., 2011; Bezerra et al., 2007; Harshbarger et al., 2017; Adams et al., 2012; Federici et al., 2009; Boskovic et al., 2013; Liao et al.; 2016; Zou et al.; 2018) Among these, piperlongumine (PL) is a natural alkaloid isolated from *Piper longum* L. characterized by the presence of two α, β- unsaturated functionalities (see **Figure 1**) and recently has been reported as a promising anticancer molecule by targeting the stress response to ROS, inducing apoptosis. (Harshbarger et al., 2017; Adams et al., 2012; Boskovic et al., 2013; Liao et al., 2016)

This molecule also represents a promising lead compound in the developing potent GSTP1 inhibitors stimulating the synthesis of a huge number of its structural analogs. (Stoddard et al., 2017; Bezerra et al., 2007; Harshbarger et al., 2017; Adams et al., 2012; Boskovic et al., 2013; Liao et al., 2016) PL acts as Michael acceptor because can undergo heteroconjugate addition with the peptide-like molecules including nucleophilic thiols of cysteine residues in irreversible or reversible fashion. From stable isotope labeling (Raj et al., 2011) the anti-cancer effects of PL were related to the promotion of reactive oxygen species (ROS) and to the reduction of GSH cellular levels. (Harshbarger et al., 2017) PL contains a trimethoxyphenyl head and two reactive olefins moieties (C2-C3 and C7-C8) that revealed to be essential for differentiating the cellular activity. (Adams et al., 2012) The C2-C3 bond is critical for toxicity, ROS elevation and protein S-glutathionylation while C7-C8 is not necessary for these activities and is believed to enhance the toxicity. (Harshbarger et al., 2017; Adams et al., 2012) This means that the two present olefins can be identified as the minimum pharmacophore of PL so that their modifications can originate analogs with different biological response. (Adams et al., 2012) Furthermore, it can act as GSTP1-cosubstrate in both displacement and addition reactions. In this case, GSH bound in the G site of GSTP1 is the target of the inhibitor. (Harshbarger et al., 2017; Adams et al., 2012) Recently, the high resolution X-ray crystal structure of GSTP1 (PDB code 5J41) in complex with PL and GSH proposed as the inhibition occurs without GSTP1 covalent modification by PL but, rather unexpectedly, PL results to be hydrolyzed to a trimethoxycinnamic acid (TMCA) deprived of the C2-C3 olefin. (Harshbarger et al., 2017) This finding does not completely fit the behavior of PL towards other cysteine-containing peptides that react with the C2-C3 reactive bond in vitro conditions. (Adams et al., 2012) Harshbarger *et al.* provided the first structural model for the interactions between PL, GSH and GSTP1. (Harshbarger et al., 2017) From this study emerged that PL acts as a prodrug. In fact, after entrance in the cell it undergoes hydrolysis giving rise to the TMCA that in turns reacts with GSH, located in the G site of GSTP1, affording the hPL:GSH conjugate as product of the addition reaction and confirming that no covalent bond formation occurs between PL and GSTP1. Although the presence in the literature of many scientific works (Bezerra et al., 2007; Harshbarger et al., 2017; Adams et al., 2012; Federici et al., 2009; Boskovic et al., 2013; Liao et al., 2016; Zou et al., 2018; Peng et al., 2015) on the piperlongumine selective inhibition of tumor growth in different types of cancers, the molecular mechanism involved in PL mediated cancer cell death remains still poorly understood. With the aim to contribute to a better knowledge, at atomistic level, of the inhibition mechanism of GSH by the

In review

This is a provisional file, not the final typeset article

hydrolyzed product of PL into the GSTP1 enzyme, a theoretical investigation in the framework of density functional theory (DFT) was undertaken. In addition, a MD simulation of initial enzyme-inhibitor (EI) complex has been also performed.

Methodology

Active site

The enzyme structure includes two identical homodimers, with a total mass of 48 kDa. The active sites are located in the interfaces between the two domains. Each active site in turn consists of two sub-sites: the *G site*, where GSH is located, in proximity to the outer side of protein surface and in direct contact with solvent molecules, and the *H site*, where the electrophilic inhibitor can be accommodated. (Harshbarger et al., 2017) The interactions between GSH and hPL with different residues of the cavity of both sites–were treated at quantum mechanical level in both QM and QM/MM calculations. In particular, the QM region includes: the Arg13 which is engaged in hydrogen bonds with the N-terminal portion of GSH and the carboxyl group of inhibitor, the Lys44 which is anchored to C-terminal part of GSH, the Tyr7 with its OH moiety oriented towards the S atom of GSH-cysteine in such a manner to establish H bond between them, and Tyr108 which is involved in π-π interaction with inhibitor aromatic ring. Finally, the QM portion, contains also the Ile104 since its crucial role in correctly orienting hPL (in *H site*) during the conjugation phase with GSH. (Harshbarger et al., 2017) Due to the closeness of active site to the protein surface, several water molecules present in the catalytic cavity were considered in the QM/MM model. Starting from the available crystallographic structure of GSTP1 by *Homo Sapiens* (Harshbarger et al., 2017) (PDB code 5J41, 1.19 Å resolution), the preparation of the models (see **Figure 2**) is described by the following procedure.

MD calculations

As first step of the work it was necessary to perform the $C8_{hPL}$-S_{GSH} bond cleavage and then to relax the enzyme:GSH:inhibitor supramolecular system at the molecular mechanics (MM) level of theory before starting the MD simulation because the used X-ray structure was related to the final product of the inhibition process. Furthermore, the presence of the inhibitor molecule of non-protein nature implied its optimization at HF/6-31G(d) level of theory in order to derive the parameters by Antechamber tool, as implemented in AMBER 16 package (AMBER version 16, 2016). Intramolecular Lennard-Jones parameters and atomic charges were obtained using, respectively, General Amber Force Field (GAFF) (Wang et al., 2004) and Restrained Electrostatic potential (RESP) method. (Bayly et al., 1993) The obtained parameters of hPL are collected in **Table S1.**

The amber ff14SB (Maier et al., 2015) force field was applied using the Xleap module and hydrogen atoms were added to the whole system. The protonation state of each amino acid has been assigned using the H++ web server. (H++, version 3.2; Anandakrishnan et al., 2012; Myers et al., 2006; Gordon et al., 2005) A rectangular box (85Åx70Åx80Å) was filled with TIP3P (Jorgensen et al., 1983) water molecules within 12.0 Å from the surface of the enzyme. The classical MD simulation was applied for 100 ps in NVT ensemble with a progressive heating phase, from 0 to 310 K. A final MD production of 20 ns was obtained in NPT ensemble (1 bar and 310 K). During the simulations, a cutoff radius for non-bonded interactions was fixed at 12 Å and Particle Mesh Ewald summation method (PME) (Ewald, 1921) and SHAKE algorithm (Ryckaert et al., 1977) were employed to constrain the motion in H-including bonds, in order to use a 2 fs integration step The root-mean-square deviation (RMSD) analysis of the whole protein and the *H* and *G* active sites residues was

In review

performed to verify the stability of the system during the MD simulation (Supporting Information, **Figure S1**). To better examine the conformational behavior of the inhibitor-protein system, a MD simulation has been also performed on the alone enzyme. The obtained root-mean-square fluctuation (RMSF) is shown in **Figure S2**. Furthermore, in order to verify conformational homogeneity for inhibitor binding modes in to the catalytic pocket, 20 structures were selected along MD simulation. (Supporting information **Figure S3**) Clustering results confirmed that the last frame, obtained at 20 ns, is a good representative configuration as to be adopted as starting configuration for creating QM cluster and QM/MM model.

QM cluster and QMMM models

The amino acids considered in the QM region (Tyr7, Arg13, Tyr103, Ile104, Lys44) were truncated as depicted in **Figure 2**. The missing hydrogens were added manually and one water molecule (lying at 3.601 Å from the GSH) was explicitly considered, being implicated in direct interaction with nucleophilic agent while the other waters are located away than 4 Å. The C atoms labelled with "*" were kept fixed during geometry optimizations, applying the locking scheme, to prevent artificial movements. (Siegbhan and Himo, 2011; Himo, 2017; Piazzetta et al., 2015) The QM cluster model was found to be adequate in the elucidation of the catalytic mechanism followed by other enzymes. (Prejanò et al., 2017; Lan and Chen, 2016; Amata et al., 2011) The obtained model consists of 136 atoms with a total charge equal to zero.

The QM/MM model was obtained applying the two layers ONIOM formalism (Svensson et al., 1996) as implemented in Gaussian09 code (Gaussian 09, Revision D.01, 2011), maintaining the same atoms mentioned in QM cluster model setup. The entire enzyme and a number of water molecules within 5 Å from the catalytic site were considered. (**Figure 2**) During the optimization, all the water molecules and residues out of 18 Å sphere from the active site were kept frozen, applying the standard procedure for single conformation PES studies. (Sousa et al., 2017) The final model contains 7811 atoms.

Technical Details

Gaussian 09 (Gaussian 09, Revision D.01, 2011) software package was used to perform calculations using B3LYP (Becke, 1993; Lee et al., 1988) hybrid functional in QM region of both used models. For S, N, H, O and C atoms, 6-31+G(d,p) basis set was used during the optimizations. Linear transit scans were performed, in order to detect stationary points along reaction coordinates. To confirm the nature of intermediates or transition states, frequencies calculation was performed at the same level of theory, for each stationary point intercepted along potential energy surface (PES). To obtain more accurate electronic energies single point calculations with 6-311+G(2d,2p) larger basis set were performed. The final energy profiles include the zero point energy (ZPE) and dispersion corrections (evaluated using the DFT-D3 procedure) (Grimme et al., 2011) and solvation energy.

The electrostatic embedding as implemented in Gaussian 09 was employed to evaluate the Coulomb interactions between MM and QM regions in all calculations. (Vreven et al., 2006) For the QM cluster calculations, single point calculations adopting the SCRF-SMD solvation model with a dielectric constant $\varepsilon = 4$, simulating the enzyme environment, was used. (Marenich et al., 2009). The same level of theory was adopted during the optimizations of species involved in hydrolysis of PL, considering the dielectric constant of 78.0, as successfully proposed in other studies. (Marino et al., 2016; Ritacco et al., 2015) NBO (NBO, version 3.1, 2001) and non covalent interaction (NCI)

In review

This is a provisional file, not the final typeset article

(NCIPLOT, version 3.0, 2011) analyses were performed on all the stationary points of the investigated PESs at both QM and QM/MM levels.

As far as the proton affinity calculations for establishing the oxygen carbonyl to be considered in the hydrolysis mechanism at acidic conditions, the proton affinities as binding energies (BE) have been estimated as indicated by the following expression:

$$BE=-(\Delta H_{hPL\text{-}H^+}-\Delta H_{hPL})$$

The BE is calculated as the difference between the enthalpy of the protonated system and that of the neutral one. In the calculations, the H^+ contribution does not appear since we evaluated the energetic difference, therefore the obtained binding energies represent the energy involved in the formation of the protonated systems.

Results and Discussion

Hydrolysis of piperlongumine

Following the recent experimental indications that demonstrate as the PL suffers hydrolysis out of the enzyme pocket, (Harshbarger et al., 2017) we firstly study this process in aqueous media. The considered reaction mechanism is illustrated in **Figure 3**. As from the experimental evidence, (Harshbarger et al., 2017) we considered the hydrolysis mechanism of PL to occur on the oxygen of the carbonyl (C6) functionality next to C7-C8 olefin, under both neutral and acidic conditions to take into account the different intracellular pH conditions, since the acid pH values are observed in cancer cells. (Wang et al., 2017; Towsend and Tew; 2003)

On the contrary, our computed BE shows that the carbonyl moiety next to C2-C3 olefin has minor proton affinity (about 4 kcal mol^{-1}) with respect to that next to C7-C8 one, indicating as under the same conditions the favored protonation site is the oxygen of C6.

The optimized geometries of the stationary points are reported in Supporting Information (**Figure S4** and **S5**), while the calculated energy profiles are depicted in **Figure 4**. As shown from **Figure 3**, we propose at acid pH a mechanism occurring in a multistep process contrary to that at neutral conditions occurring in only one step. In both cases the product is the hPL, while the leaving group is 1,2,5,6-tetra-hydro-pyridin-2-ol (t-PyrOH) neutral and protonated, respectively. For clarity, in the text the remaining double bond in the hPL upon hydrolysis will retain the same numeration than in PL, (C7-C8). The processes are exothermic although at pH acid the exoergonicity is more pronounced. (**Figure 4**) From our results, the acidic hydrolysis is strongly favored as suggested by lower activation barriers (by about 10 kcal mol^{-1}) than that found at neutral pH (see **Figure 4**). The calculated barrier in acidic environment well agrees with those characterizing other anticancer molecules acting as prodrug. (Alberto and Russo, 2011; Marino et al., 2017; Ritacco et al., 2015) Once the hydrolyzed product is formed, the process of GSH-conjugation favored by GSTP1 starts through the attack to the C7-C8 double bond.

GSTP1 inhibition

To underline the role of GSTP1 during the inhibition process by hPL, we have considered, at both QM and QM/MM levels, two different reaction mechanisms (**A**, and **B**) as presented in **Figure 5**. In particular: (**A**) describes the nucleophile addition to the double bond of inhibitor by -SH group of GSH without involving any amino acid residue while path (**B**) takes into account the participation of

In review

the Tyr7 residue in the formation of the covalent adduct. In all the cases, the inhibition reaction occurs in a one-step process by the Michael addition of the thiol from GSH at the C7-C8 olefin of hPL. In all the considered mechanisms, the starting species is the ternary enzyme-hydrolyzed inhibitor-GSH complex (**EI**) obtained after the geometry optimization of the frame isolated by the previous MD cycle at 20 ns. From **Figure 6**, that illustrates the superposition of the crystallographic structure with the last MD snapshot, it is possible to note that hPL interacts with the binding cleft of the *H site* and no water molecules are close to the reaction center, in agreement with the hydrophobic nature of site (Tyr7, Tyr108 and Ile104 residues). As expected, the crystallographic pose (obtained with the reaction product) deviates in this moiety (see **Figure 6**). At the contrary, the GSH region is well superimposed confirming that this molecule is well positioned with a correct orientation of the thiol.

The energy profiles obtained employing QM and QM/MM tools for the two considered mechanisms, are reported in **Figure 7**. The reported QMMM energy values do not include the entropic contribution. In order to quantify this the Grimme procedure has been employed. (Grimme, 2012) Results (see **Table S2**) evidence that the TΔS terms slightly affect the previously obtained energy values. QM/MM structures of the initial complex (**EI**), the final S-conjugated product (**P**) and that of the transition states are reported in **Figures 8**. All the QM optimized geometries are given in Supporting Material (**Figure S6**).

In **EI** the carboxylate moiety of the hPL is oriented in such a way to establish hydrogen bonds with Tyr7 (1.599 Å) and SH group of GSH (2.138 Å). Furthermore, van der Waals and hydrophobic interactions, such as those between the inhibitor and the Ile104 and Arg14 residues (see **Figure 8**) contribute to optimally accommodate the inhibitor into the H binding site. In fact, now the key reacting atoms, C8 of the hPL and SH nucleophile species, are placed in suitable way (at 4.271 Å) to allow the deactivation reaction. In path **A** the intercepted transition state (**TSA**) represents a four-centered structure where the sulphur addition to the C8 (1.880 Å) and the C7-H bond formation (1.467 Å) simultaneously occur. The corresponding frequency analysis confirms a first-order saddle point with an imaginary frequency (1510*i* cm^{-1}) which corresponds to a vibrational mode involving a strong C7−H coupling and a relatively weaker C8−S one. The C8-S bond is already established and the forming C-H one can be evinced by the elongation of the C7-C8 bond (1.544 Å). This barrier results to be 64.4 kcal mol^{-1} at QM/MM level and 70.9 kcal mol^{-1} at QM one. Both values are very close to that computed for the reaction unassisted by the catalyst (76.0 kcal mol^{-1}, see **Figure 7**). The resulting product (**P**), shown in **Figure 8,** evidences that the C-S bond is formed (1.818 Å) and the C8-C7 is elongated (1.532 Å) confirming the occurred sp^3 hybridization of the two involved carbon atoms. The exothermicity is evaluated to be 10.8 kcal mol^{-1} (5.8 kcal mol^{-1} in the QM cluster). The mechanism **B** (**Figure 5**) involves the participation of the Tyr7 residue. In **Figure 8** is reported the optimized structure of the **TSB** connecting the **EI** and the covalent final complex (**P**). The nucleophilic attack to C8 occurs by GSH-thiolate (1.914 Å) since the hydrogen of the S-H_{GSH} group (2.019 Å) has been delivered to oxygen (O_{Tyr}) of the side chain of Tyr7 (1.090 Å). In fact, the OH group of Tyr7, oriented via hydrogen bonding to carboxylate moiety of the inhibitor (1.599 Å), in the TS becomes 1.310 Å and points towards C7 atom for delivering its hydrogen atom (O-H and H-C7 distances are found to be 1.310 Å and 1.174 Å, respectively) while the C7-C8 bond is elongated (1.535 Å) (see **Figure 8)**. The TS located along the mechanism **B** lies at 21.5 kcal mol^{-1} (QM) and 19.8 kcal mol^{-1} (QM/MM) above the **EI**. Both values are very close to the available experimental one (23.8 kcal mol^{-1}) concerning the C-S bond formation. (Huskey et al., 1991) The superposition of our optimized glutathionil-conjugated product **P** with the corresponding crystallographic structure (Harshbarger et al., 2017) (see **Figure 9**) reveals a good RMSD value in both GSH and *H site* regions. The exothermicity 10.8 kcal mol^{-1} means that the reverse reaction can be accessible but

In review

This is a provisional file, not the final typeset article

much slower also for the high barrier required in the reverse process **P→ EI** (30.6 kcal mol^{-1}). **TSB** evidence as the formation of the S-C8 bond is strictly related to the deprotonation of the SH moiety of GSH at the expense of the Tyr7 acting as proton shuttle with a consequent reduction of barrier (19.8 kcal mol^{-1})

This is in agreement with previous works on other GST enzymes (Dourado et al., 2008; Zheng and Ornstein, 1997; Angelucci et al., 2005) revealing the importance of the acidic properties of a Tyr during the catalysis of glutathione-S-Transferase. Furthermore, our findings corroborate the hypothesis advanced by the previous structural analysis (Harshbarger et al., 2017) revealing as no covalent bond formation between hPL and GSPT1 was observed and that PL acts as prodrug. With the aim to evaluate the nature of the interactions present inside the catalytic pocket during the process, in **Figure 10** we have reported the density of isosurfaces arising from NCI analysis, indicating the different contributions of the residues retained in the QM region.

In every characterized stationary point, it can be noted the salt bridges occurring between the side-chains of Lys44 and Arg13 with the carboxyl moiety of carboxyl- and amino-terminal of GSH (blue region indicates strong attractive interactions while the red isosurfaces account for the repulsive interactions related to the center of π systems of Tyr7 and Tyr108 and the inhibitor molecule, as usually for aromatic systems strong nonbonded overlap is indicated). (Johnson et al., 2010) Further information arises from the green regions indicative of the van der Waals forces characterizing the cavity containing the inhibitor molecule and identified by the hydrophobic residues Tyr108 and Ile104. It is interesting to underline as the interaction involving the Ile104 becomes more intense as the reaction proceeds. At the contrary no relevant contributions arise from the NBO analysis (see **Table S3**) except for a little bit increased nucleophile nature of the sulphur atom of GSH in the enzyme and a decreased negative charge on the C7-C8 bond of the hPL species with the respect to the corresponding values obtained for the process unassisted by enzyme. In the **TSB** species, a more attractive interaction appears in proximity of the region interested in the chemical events (circled in red in **Figure 10**) symptomatic of the occurring S-C8 bond formation.

Conclusions

This study focuses on the inhibition mechanism of the glutathione-S-transferase Pi 1 by the hydrolyzed product of piperlongumine. We propose the mechanism following the most recent experimental evidences taking into account in particular the role of Tyr7 on the complex formation between the glutathione and the inhibitor inside the catalytic pocket of enzyme.

The hydrolysis of PL for giving hPL has been considered in neutral and acid conditions. The last one provided the better energetic path.

The agreement between cluster QM and the more computational demanding hybrid QM/MM methods is quite good. Structural and energetic computed properties are in line with the available experimental data.

The lowest energy reaction mechanism for reaction of hPL with GSH corresponds to that in which the Tyr7 residue is involved in the inhibition reaction deprotonating the GSH and donates the proton, in a concerted fashion, to the C7 substrate atom. The computed barrier heights result 19.8 and 21.5 kcal mol^{-1} in both QM/MM and QM models respectively. Both computations clearly indicate the same reaction mechanism by TSB as the preferred one with difference in the barrier eight is of only 1.7 kcal $mol^{-1}$and propose an exergonic reaction.

In review

We hope that the obtained new insights on the reaction mechanism of human GSTP1 inhibition with natural piperlongumine substrate can be useful in the design of new selective and more potent inhibitors.

Reference

Adams, D. J., Dai, M. G., Pellegrino, G., Wagner, B. K., Stern, A. M., Shamji, A. F., and Schreiber, S. L. (2012) Synthesis, cellular evaluation, and mechanism of action of piperlongumine analogs. *PNAS* 109, 15115-15120. doi:10.1073/pnas.1212802109

Adler, V., Yin, Z., Fuchs, S. Y., Benezra, M., Rosario, L., Tew, K. D., Pincus, M. R., Sardana, M., Henderson, C. J., Wolf, C. R., Davis, R. J., and Ronai, Z, (1999) Regulation of JNK signaling by GSTp. *EMBO J.* 18, 1321–1334. doi:10.1093/emboj/18.5.1321

Alberto, M. E., and Russo, N. (2011) Methionine ligand selectively promotes monofunctional adducts between trans-EE platinum anticancer drug and guanine DNA base. *Chem. Commun.* 47, 887–889. doi:10.1039/C0CC03605F

Amata, O., Marino, T., Russo, N., and Toscano, M. (2011) Catalytic activity of a ζ-class zinc and cadmium containing carbonic anhydrase. Compared work mechanisms. *Phys. Chem. Chem. Phys.* 13, 3468-3477. doi:10.1039/C0CP01053G

AMBER 16 (2016), University of California, San Francisco

Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 40(W1), W537-541. doi:10.1093/nar/gks375

Angelucci, F., Baiocco, P., Brunori, M., Gourlay, L., Morea, and V., Bellelli, A. (2005) Insights into the Catalytic Mechanism of Glutathione S-Transferase: The Lesson from Schistosoma haematobium. *Structure* 13, 1241–1246. doi: 10.1016/j.str.2005.06.007

Armstrong, R. N. (1997) Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol.* 10, 2–18. doi:10.1021/tx960072x

Ban, N., Takahashi, Y., Takayama, T., Kura, T., Katahira, T., Sakamaki, S., and Niitsu, Y. (1996) Transfection of Glutathione S-Transferase (GST)-π Antisense Complementary DNA Increases the Sensitivity of a Colon Cancer Cell Line to Adriamycin, Cisplatin, Melphalan, and Etoposide. *Cancer Res.* 56, 3577–3582.

Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* 97(40), 10269-10280. doi:10.1021/j100142a004

Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 98, 5648-5652. doi: 10.1063/1.464913

Bezerra, D. P., Militão, G. C., de Castro, F. O., Pessoa, C., de Moraes, M. O., Silveira, E. R., Lima, M. A., Elmiro, F. J., and Costa-Lotufo, L. V. (2007) Piplartine induces inhibition of leukemia cell

In review

This is a provisional file, not the final typeset article

proliferation triggering both apoptosis and necrosis pathways. *Toxicol. In Vitro*, 21, 1-8. doi:10.1016/j.tiv.2006.07.007

Boskovic, Z. V., Hussain, M. M., Drew, J. A., Mingji, D., and Schreiber, S. L. (2013) Synthesis of piperlogs and analysis of their effects on cells. *Tetrahedron* 69, 7559-7567. doi:10.1016/j.tet.2013.05.080

Broxterman, H. J., Giaccone, G., and Lankelma, J. (1995) Multidrug resistance proteins and other drug transport-related resistance to natural product agents. *Curr. Opin. Oncol.* 7, 532-540. doi:10.1097/00001622-199511000-0001

Dirr, H., Reinemer, P., Huber, R. (1994) X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. *Eur. J. Biochem.* 220, 645-661. doi:10.1111/j.1432-1033.1994.tb18666.x

Dourado, D. F. A. R., Fernandes, P. A., Mannervik, B., and Ramos, M. J. (2008) Glutathione Transferase: New Model for Glutathione Activation. *Chem. Eur. J.* 14, 9591-9598. doi:10.1002/chem.200800946.

Ewald, P. P.(1921) *Ann. Phys-Berlin* 64(3), 253-287.

Federici, L., Lo Sterzo, C., Pezzola, S., Di Matteo, A., Scaloni, F., Federici, G., and Caccuri, A. M. (2009) Structural basis for the binding of the anticancer compound 6-(7-nitro-2,1,3-benzoxadiazol-4-ylthio)hexanol to human glutathione s-transferases. *Cancer Res.* 69, 8025-8034. doi:10.1158/0008-5472.CAN-09-1314

Gaussian 09, Revision D.01 (2011), *Gaussian*, Inc., Wallingford CT.

Gordon, J. C., Myers, J. B., Folta, T. M., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H++: a server for estimating p Ka s and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 33,W368-371. doi: 10.1093/nar/gki464

Green, J. A., Robertson, L. J., and Clark, A. H. (1993) S-transferase expression in benign and malignant ovarian tumours. *Br. J. Cancer* 68, 235–239.

Grignon, D. J., Abdel-Malak, M., Mertens, W. C., Sakr, W. A., and Shepherd, R. R. (1994) Glutathione S-transferase expression in renal cell carcinoma: a new marker of differentiation. *Mod. Pathol.* 7, 186–189.

Grimme, S. , Ehrlich, S., and Goerigk, L.(2011) Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* 32, 1456-1465. doi:10.1002/jcc.21759

Grimme, S. (2012), Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.*18, 9955-9964. doi: 10.1002/chem.201200497

H++, vesion 3.2 (2016). http://biophysics.cs.vt.edu/H++

Harshbarger, W., Gondi, S., Ficarro, S. B., Hunter, J., Udayakumar, D., Gurbani, D., Singer, W. D., Liu, Y., Li, L., Marto, J. A., and Westover, K. D.(2017) Structural and Biochemical Analyses Reveal

In review

the Mechanism of Glutathione S-Transferase Pi 1 Inhibition by the Anti-cancer Compound Piperlongumine. *J. Biol. Chem.* 292, 112-120. doi:10.1074/jbc.M116.750299

Hayes, J. D., Flanagan, J. U., and Jowsey, I. R. (2005) Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* 45, 51-88. doi: 10.1146/annurev.pharmtox.45.120403.095857

Himo, F. (2017) Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.*139, 6780-6786. doi:10.1021/jacs.7b02671

Huskey, S.-E. W., Phillip Huskey, W., and Lu, A. Y. H. (1991) Contributions of thiolate "desolvation" to catalysis by glutathione S-transferase isozymes 1-1 and 2-2: evidence from kinetic solvent isotope effects. *J. Am. Chem. Soc.* 113, 2283-2290. doi:10.1021/ja00006a056

Inoue, T., Ishida, T., Sugio, K., Maehara, Y., and Sugimachi, K. (1995) Glutathione S transferase Pi is a powerful indicator in chemotherapy of human lung squamous-cell carcinoma. *Respiration* 62, 223–227. doi:10.1159/000196451

Johnson, E. R., Keinan, S., Mori-Sanchez, P., Contreras-Garcia, J., Cohen, A. J., and Yang, W. (2010) Revealing noncovalent interactions. *J. Am. Chem. Soc.*, 132, 6498-6506. doi: 10.1021/ja100936w

Jorgensen, W., Chandrasekhar, J., Madura, J., and Klein, M. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926-935. doi:10.1063/1.445869

Karpusas, M , Axarli, I., Chiniadis, L., Papakyriakou, A., Bethanis, K., Scopelitou, K., Clonis, Y. D., and Labrou, N. E. (2013) The interaction of the chemotherapeutic drug chlorambucil with human glutathione transferase A1-1: kinetic and structural analysis. *PLoS ONE* 8(2), e56337. doi:10.1371/journal.pone.0056337

Katagiri, A., Tomita, Y., Nishiyama, T., Kimura, M., and Sato, S. (1993) Immunohistochemical detection of P-glycoprotein and GSTP1-1 in testis cancer. *Br. J. Cancer*. 68, 125–129. doi:10.1038/bjc.1993.299

Lan, C.-L., and Chen, S.-L. (2016) The Decarboxylation of α,β-Unsaturated Acid Catalyzed by Prenylated FMN-Dependent Ferulic Acid Decarboxylase and the Enzyme Inhibition. *J. Org. Chem.* 81, 9289-9295. doi:10.1021/acs.joc.6b01872

Lee, C. T. ; Yang, W T., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37, 785-789. doi:10.1103/PhysRevB.37.785

Liao, Y., Niu, X., Chen, B., Edwards, H., Xu, L., Xie, C., Lin, H., Polin, L., Taub, J. W., Ge, Y., and Qin, Z. (2016) Synthesis and Antileukemic Activities of Piperlongumine and HDAC Inhibitor Hybrids against Acute Myeloid Leukemia Cells. *J. Med. Chem.* 59, 7974-7990. doi:10.1021/acs.jmedchem.6b00772

Lo, H. W., and Ali-Osman, F. (2007) Genetic polymorphism and function of glutathione S-transferases in tumor drug resistance. *Curr. Opin. Pharmacol.* 7, 367-374. doi:10.1016/j.coph.2007.06.009

In review

This is a provisional file, not the final typeset article

Marenich, A. V. , Cramer, C. J., and Truhlar, D. G. (2009) Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* 113(8), 6378-6396. doi:10.1021/jp810292n

Marino, T., Parise, A., and Russo, N. (2016) The role of arsenic in the hydrolysis and DNA metalation processes in an arsenous acid–platinum(II) anticancer complex. *Phys. Chem. Chem. Phys.* 19, 1328—1334. doi:10.1039/C6CP06179F

McIlwain, C. C., Townsend, D. M., and Tew, K. D. (2006) Glutathione S-transferase polymorphisms: cancer incidence and therapy. *Oncogene* 25, 1639-1648. doi:10.1038/sj.onc.1209373

Mejerman, I., Beijnen, J. H., and Schellens, J. H. (2008) Combined action and regulation of phase II enzymes and multidrug resistance proteins in multidrug resistance in cancer. *Cancer Treat. Rev.* 34, 505-520. doi:10.1016/j.ctrv.2008.03.002

Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* 63, 928-938. doi:10.1002/prot.20922

NBO, version 3.1 (2001).

NCIPLOT, version 3.0(2011). Download: http://www.lct.jussieu.fr/pagesperso/contrera/nciplot.html

Oakley, A. J. (2005) Glutathione transferases: new functions. *Curr. Opin. Struct. Biol.* 15, 716-723. doi:10.1016/j.sbi.2005.10.005

Oakley, A. J., Rossjohn, J., Lo Bello, M., Caccuri, A. M., Federici, G., and Parker, M. W. (1997) Multifunctional role of Tyr 108 in the catalytic mechanism of human glutathione transferase P1-1. Crystallographic and kinetic studies on the Y108F mutant enzyme. *Biochemistry* 36, 576-585. doi:10.1021/bi962813z

Okuyama, T., Maehara, Y., Endo, K., Baba, H., Adachi, Y., Kuwano, M., and Sugimachi, K.(1994) Expression of glutathione s-transferasepi and sensitivity of human gastric cancer cells to cisplatin. *Cancer* 74, 1230–1236. doi:10.1002/1097-0142(19940815)74:4<1230::AID-CNCR2820740409>3.0.CO;2-0

Pei, S., Minhajuddin, M., Callahan, K. P., Balys, M., Ashton, J. M., Neering, S. J., Lagadinou, E. D., Corbett, C., Ye, H., Liesveld, J. L., O'Dwyer, K. M., Li, Z., Shi, L., Greninger , P., Settleman, J., Benes, C., Hagen, F. K., Munger, J., Crooks, P. A., Becker, M. W., Jordan, C. T. (2013) Targeting Aberrant Glutathione Metabolism to Eradicate Human Acute Myelogenous Leukemia Cells. *J. Biol. Chem*, 288, 33542- 33558. doi:10.1074/jbc.M113.511170

Peng, S., Zhang, B., Meng, X., Yao, Y., Fang, J. (2015) Synthesis of Piperlongumine Analogues and Discovery of Nuclear Factor Erythroid 2-Related Factor 2 (Nrf2) Activators as Potential Neuroprotective Agents. *J. Med. Chem.* 58(13), 5242–5255. doi:10.1021/acs.jmedchem.5b00410

Perperopoulou, F., Pouliou, F., and Labrou, N. E. (2018) Recent advances in protein engineering and biotechnological applications of glutathione transferases. *Crit. Rev. Biotechnol.* 38, 511-528. doi:10.1080/07388551.2017.1375890

In review

Piazzetta, P., Marino, T., and Russo, N., and Salahub, D. R. (2015) Direct Hydrogenation of Carbon Dioxide by an Artificial Reductase Obtained by Substituting Rhodium for Zinc in the Carbonic Anhydrase Catalytic Center. A Mechanistic Study. *ACS Catal.* 5, 5397−5409. doi:10.1021/acscatal.5b00185

Prejanò, M., Marino, T., and Russo, N. (2017) How Can Methanol Dehydrogenase from Methylacidiphilum fumariolicum Work with the Alien Ce^{III} Ion in the Active Center? A Theoretical Study. *Chem. Eur. J.*23, 8652-8657. doi:10.1002/chem.201700381

Raj, L., Ide, T., Gurkar, A. U., Foley, M., Schenone, M., Li, X., Tolliday, N. J., Golub, T. R., Carr, S. A., Shamji, A. F., Stern, A. M., Mandinova, A., Schreiber, S. L., and Lee, S. W. (2011) Selective killing of cancer cells by a small molecule targeting the stress response to ROS. *Nature* 475, 231-234. doi:10.1038/nature10167.

Ritacco, I., Russo, N., and Sicilia, E. (2015) DFT Investigation of the Mechanism of Action of Organoiridium(III) Complexes As Anticancer Agents. *Inorg. Chem.* 54, 10801–10810. doi:10.1021/acs.inorgchem.5b01832

Ruiz-Gomez, M. J., Souviron, A., Martinez-Morillo, M., and Gil, L. (2000) P-glycoprotein, glutathione and glutathione S-transferase increase in a colon carcinoma cell line by colchicine. *J. Physiol. Biochem.* 56, 307–312. doi:10.1007/BF03179798

Ryckaert, J-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes *J. Comput. Phys.* 23(3), 327-341. doi:10.1016/0021-9991(77)90098-5

Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. (2001) Structures, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* 360, 1–16. doi:10.1042/bj3600001

Siegbahn, P. E. M. , and Himo, F. (2011) The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdiscip. Rev: Comput. Mol. Sci.*1, 323-336. doi:10.1002/wcms.13

Sousa, S. F. , M. Ribeiro, A. J., Neves, R. R. P., Brás, N. F., Cerqueira, N. M. F. S. A., Fernandes, P. A., M. J. Ramos, M. J. (2017) Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *WIREs Comput. Mol. Sci.* 7, e1281-1311. doi:10.1002/wcms.1281

Stoddard, E. G., Killinger, B. J., Nair, R. N., Sadler, N. C., Volk, R. F., Purvine, S. O., Shukla, A. K. Smith, J. N., and Wright, A. T. (2017) Activity-Based Probes for Isoenzyme- and Site-Specific Functional Characterization of Glutathione S-Transferases. *J. Am. Chem. Soc.* 139, 16032−16035. doi:10.1021/jacs.7b07378

Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels−Alder Reactions and $Pt(P(t\text{-}Bu)_3)_2 + H_2$ Oxidative Addition. *J. Phys. Chem.* 100, 19357-19363. doi:10.1021/jp962071j

In review

This is a provisional file, not the final typeset article

Tew, K. D. (2007) Redox in redux: Emergent roles for glutathione S-transferase P (GSTP) in regulation of cell signaling and S-glutathionylation. *Biochem. Pharmacol.* 73, 1257–1269. doi:10.1016/j.bcp.2006.09.027

Townsend, D. M., and Tew, K. D. (2003) The role of glutathione-S-transferase in anti-cancer drug resistance. *Oncogene* 22, 7369-7375. doi:10.1038/sj.onc.1206940

Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, J. A. Jr., Morokuma, K, and Frisch, M. J. (2006) Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *J. Chem. Theory Comput.* 2, 815-826. doi:10.1021/ct050289g

Wang, H. B., Jin, X. L. , Zheng, J. F., Wang, F., Dai, F., and Zhou, B. (2017) Developing piperlongumine-directed glutathione S-transferase inhibitors by an electrophilicity-based strategy. *Eur. J. Med. Chem.* 126, 517-525. doi:10.1016/j.ejmech.2016.11.034

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comp. Chem.* 25, 1157-1173. doi:10.1002/jcc.20035

Wang, T., Arifoglu, P., Ronai, Z., and Tew, K. D. (2001) Glutathione S-transferase P1–1 (GSTP1–1) Inhibits c-Jun N-terminal Kinase (JNK1) Signaling through Interaction with the C Terminus. *J. Biol. Chem.* 276, 20999–21003. doi:10.1074/jbc.M101355200

Wilce, M. C., and Parker, M. W. (1994) Structure and function of glutathione S-transferases. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1205, 1–18. doi:10.1016/0167-4838(94)90086-8

Zhang, L., Xiao, Y., and Priddy, R. (1994) Increase in placental glutathione S-transferase in human oral epithelial dysplastic lesions and squamous cell carcinomas. *J. Oral. Pathol. Med.* 23, 75–79. doi:10.1111/j.1600-0714.1994.tb00260.x

Zheng, Y.-J., and Ornstein, R: L. (1997) Role of Active Site Tyrosine in Glutathione S-Transferase: Insights from a Theoretical Study on Model Systems. *J. Am. Chem. Soc.* **1997**, *119*, 1523-1528. doi:10.1021/ja961667h

Zou, Y., Zhao, D., Yan, C., Ji, Y., Liu, J., Xu, J., Lai, Y., Tian, J., Zhang, Y., and Huang, Z. (2018) Novel Ligustrazine-Based Analogs of Piperlongumine Potently Suppress Proliferation and Metastasis of Colorectal Cancer Cells in Vitro and in Vivo. *J. Med. Chem.*, 61, 1821–1832. doi:10.1021/acs.jmedchem.7b01096

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

MP, TM, and NR have analyzed the results, edit and reviewed equally the manuscript. MP, TM, and NR approved it for publications.

Acknowledgments

In review

Financial support from the Università degli Studi della Calabria-Dipartimento di Chimica e Tecnologie Chimiche (CTC) is acknowledged.

Supplementary Material

Supplementary Material should be uploaded separately on submission, if there are Supplementary Figures, please include the caption in the same file as the figure. Supplementary Material templates can be found in the Frontiers Word Templates file.

Figure 1. Chemical Species involved in the inhibition reaction of GSTP1.

Figure 2. The two models used. The QM/MM model (a) includes the whole protein with the water molecules, the red box defines the QM region used in both cluster (b) and QM/MM calculations. The amino acid residues of the QM portion are shown with ball and sticks representation.

Figure 3. Hydrolysis mechanism of PL in neutral (bold line) and acidic conditions (dashed line).

Figure 4. B3LYP-D3/6-311+G(2d,2p)//B3LYP/6-31+G(d,p) (SMD,ε=78) energy profiles of PL hydrolysis in neutral conditions (blue line) and acid conditions (red line). The final energies contain ZPE and D3 corrections.

Figure 5. Proposed mechanisms for GSTP1 inhibition by hPL.

Figure 6. Superposition of the ternary adduct (**EI**) from MD (violet) with the X-ray structure (yellow) related to the **P** S-conjugate product.

Figure 7. /B3LYP-D3/6-311+G(2d,2p)//B3LYP/6-31+G(d,p) (SMD,ε=4) (top) and B3LYP-D3/6-311+G(2d,2p):ff99SB//B3LYP/6-31+G(d,p):ff99SB (bottom) energy profiles of GSTP1 inhibition process by hPL for **A**, **B** mechanisms. In the black window are depicted the energy profile related to the reaction unassisted by the enzyme. The final energies contain ZPE and D3 corrections.

Figure 8. B3LYP/6-31+G(d,p):ff99SB optimized geometries of **EI**, **P**, **TSA**, **TSB**. For the transition states the imaginary frequencies are reported. .

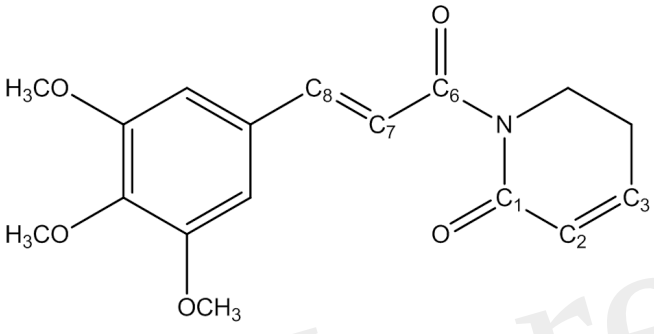
Figure 9. Superposition of the ONIOM optimized structure of the final hPL:GSH conjugated product (violet) with its corresponding crystallographic structure characterized (yellow).

Figure 10. Nonbonding interaction plot calculated for the stationary points at B3LYP-D3/6-311+G(2d,2p)//B3LYP/6-31+G(d,p) level for the **B** mechanism. The red circle defines the portion where the bonds breaking and formation occur.

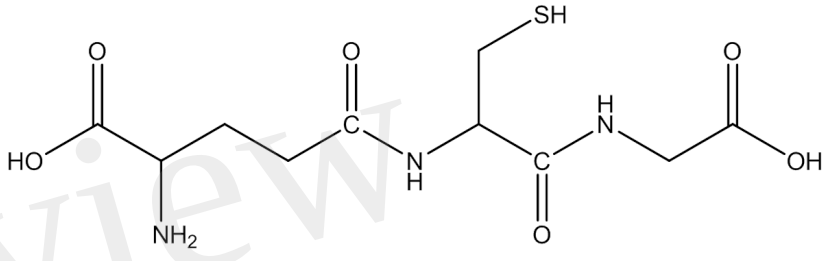
In review

This is a provisional file, not the final typeset article

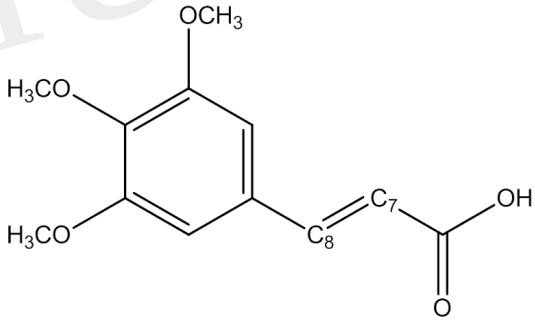
Figure 1.TIF



Piperlongumine (PL)



Glutathione (GSH)

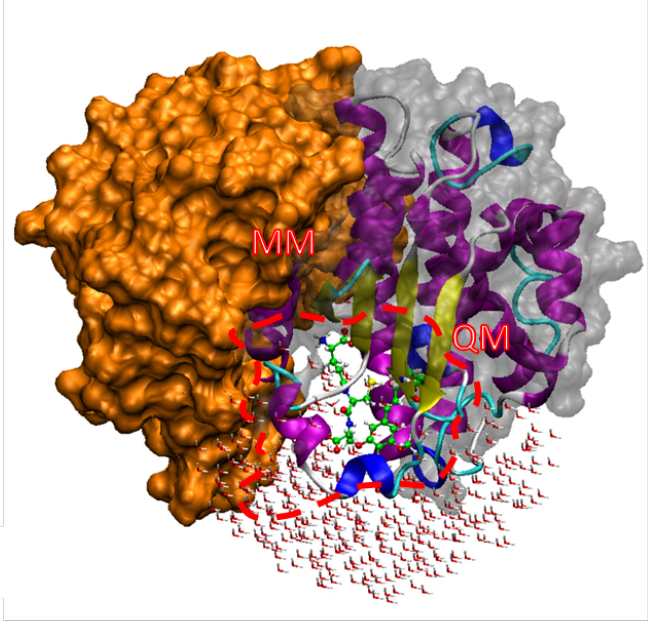


Trimethoxycinnamic acid (TMCA or hPL)

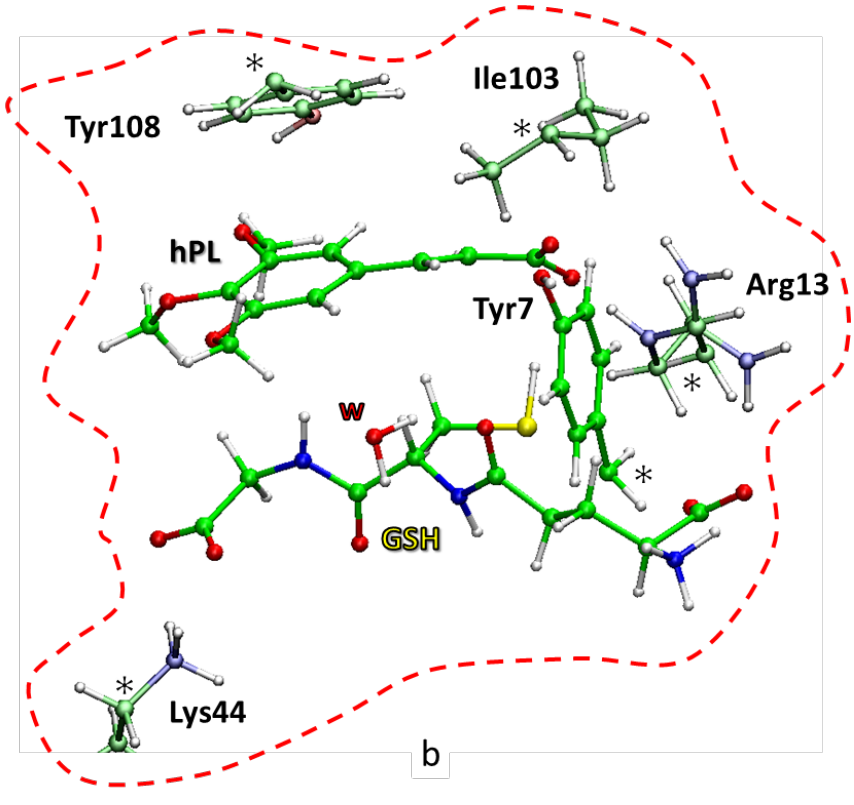
In review

Figure 2.TIF

In review

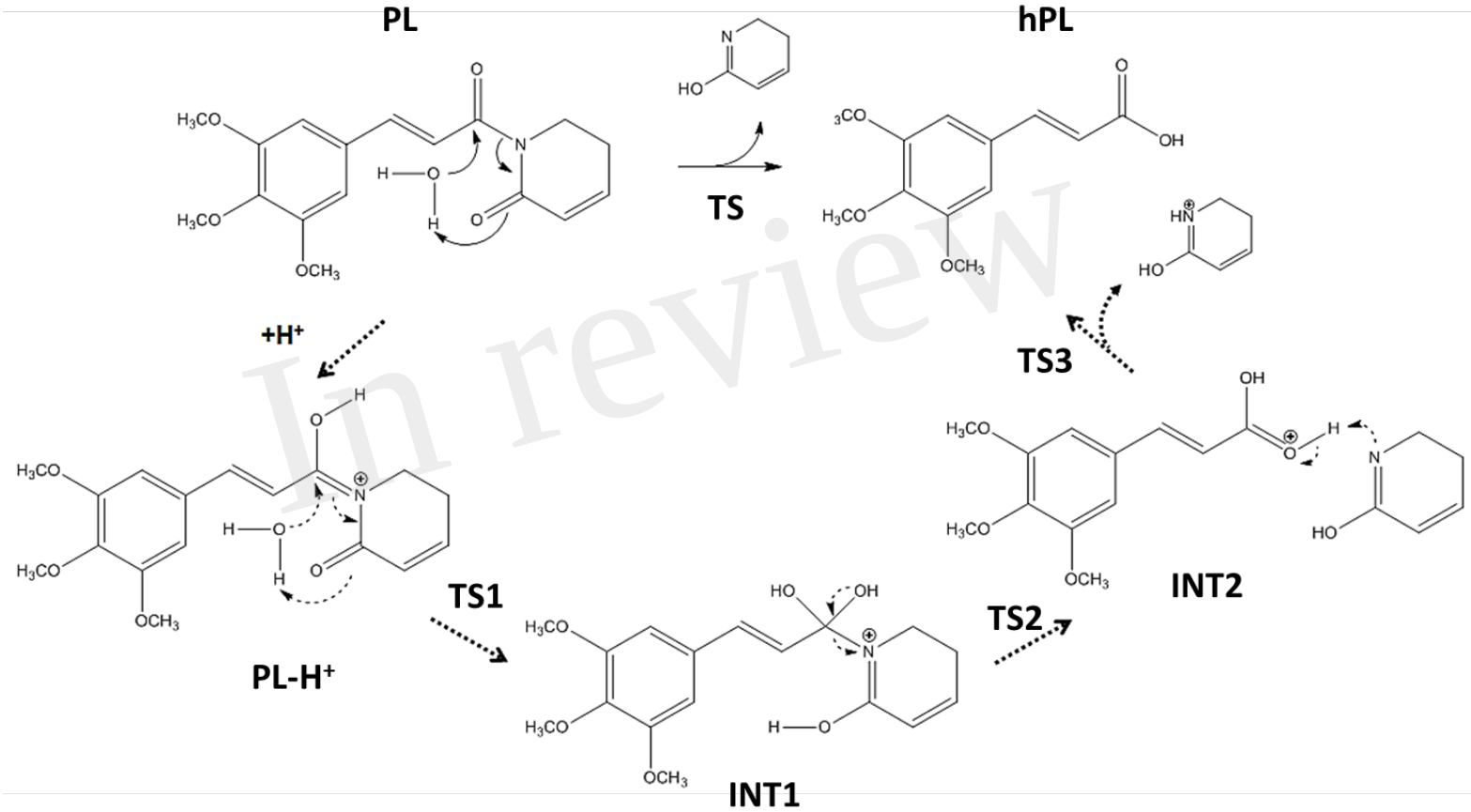


a



b

Figure 3.TIF



In review

Figure 4.TIF

In review

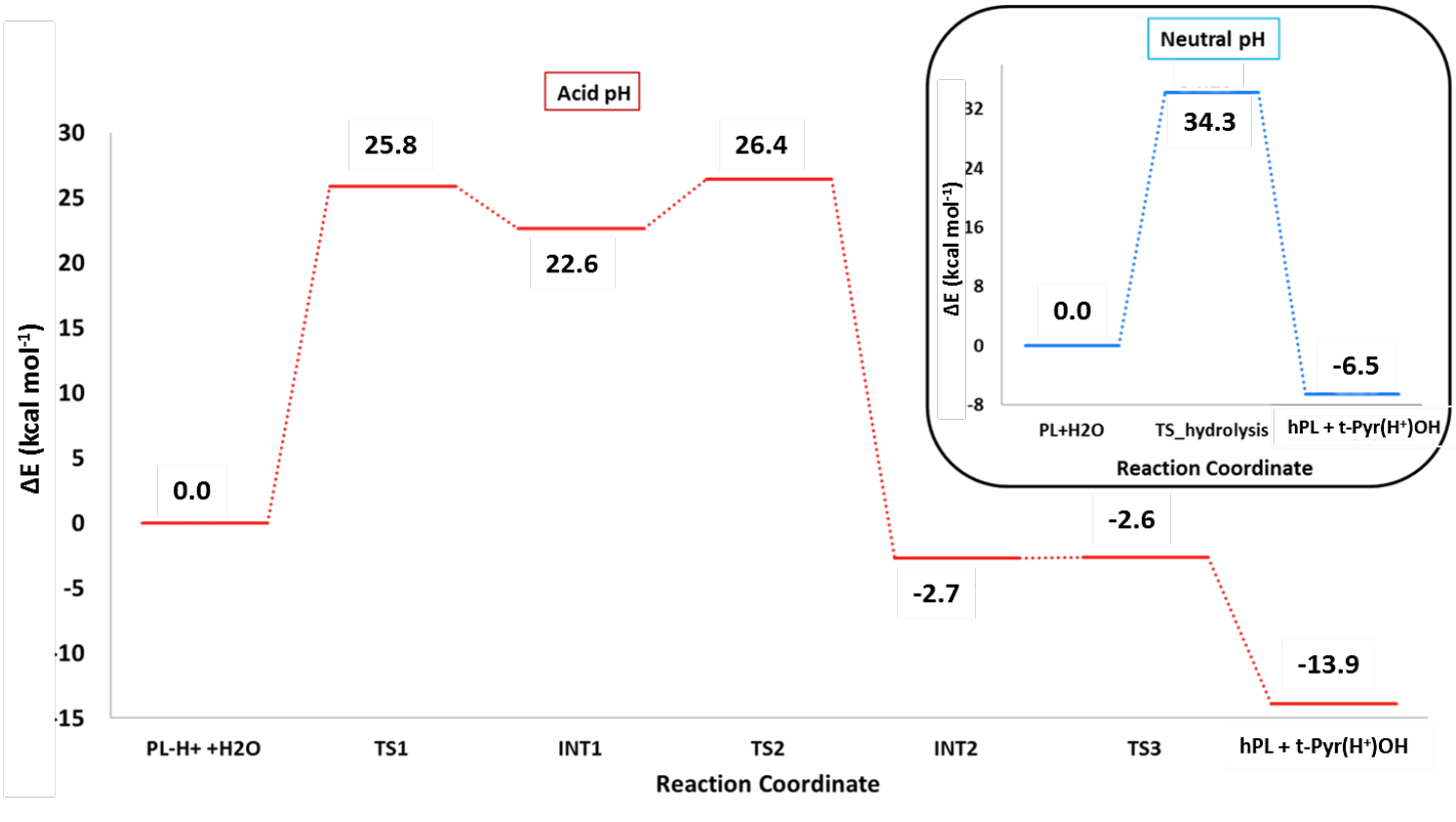


Figure 5.TIF

In review

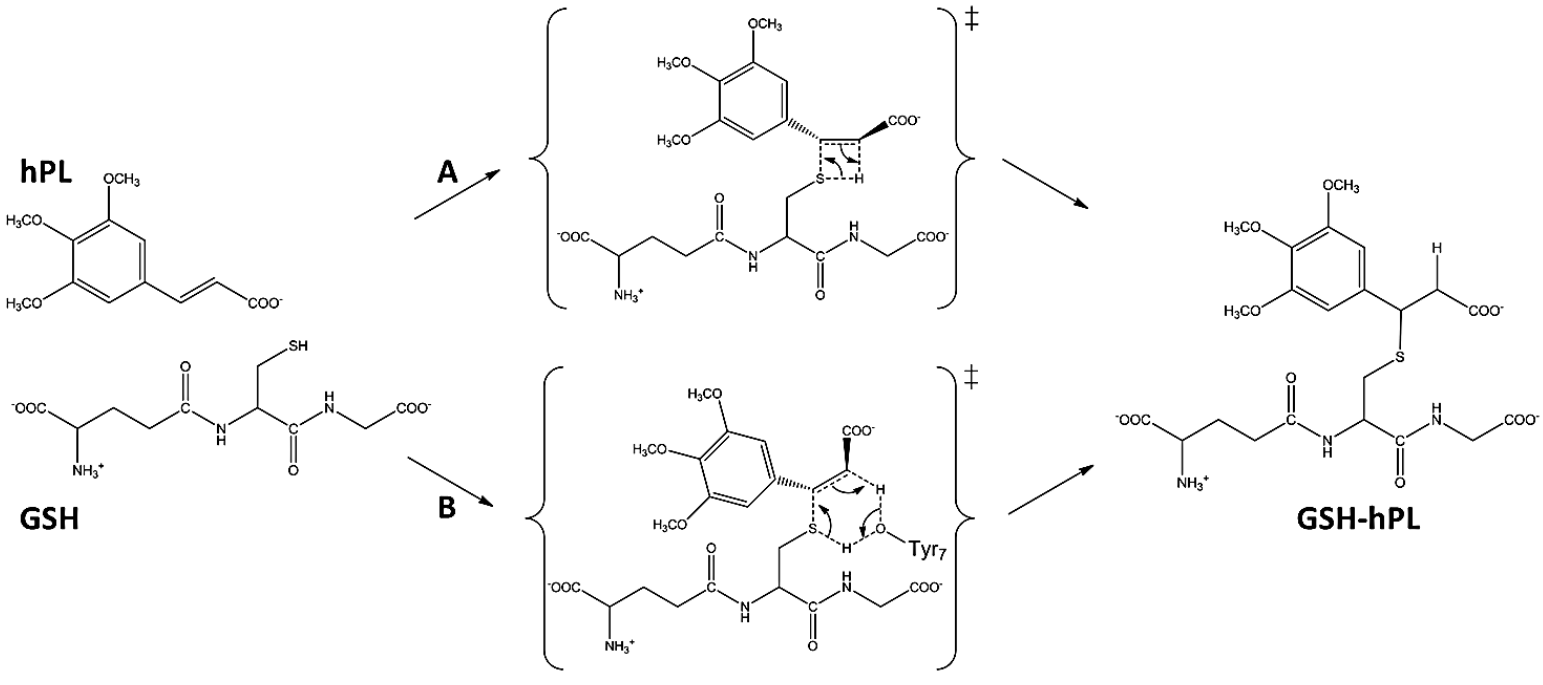


Figure 6.TIF

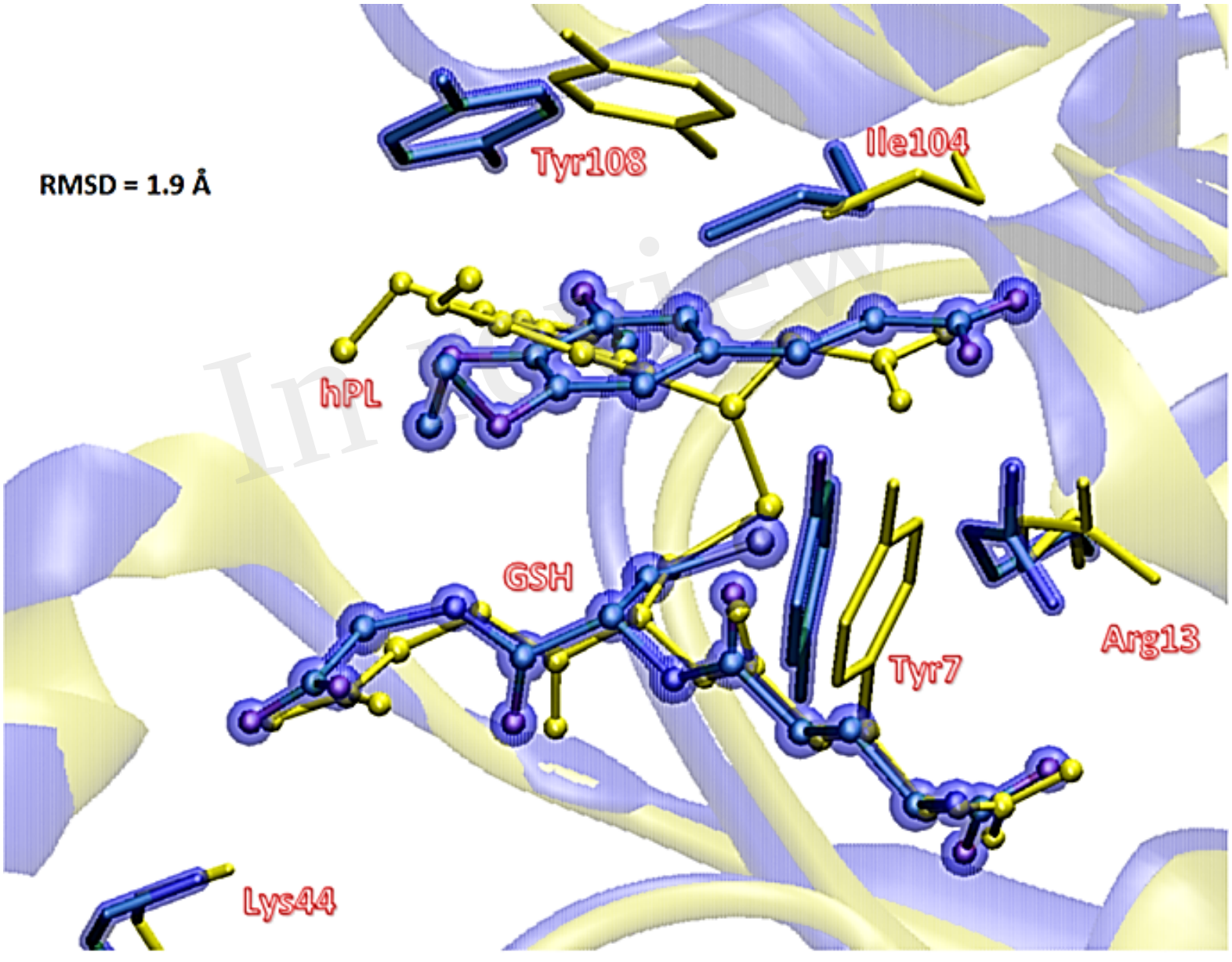
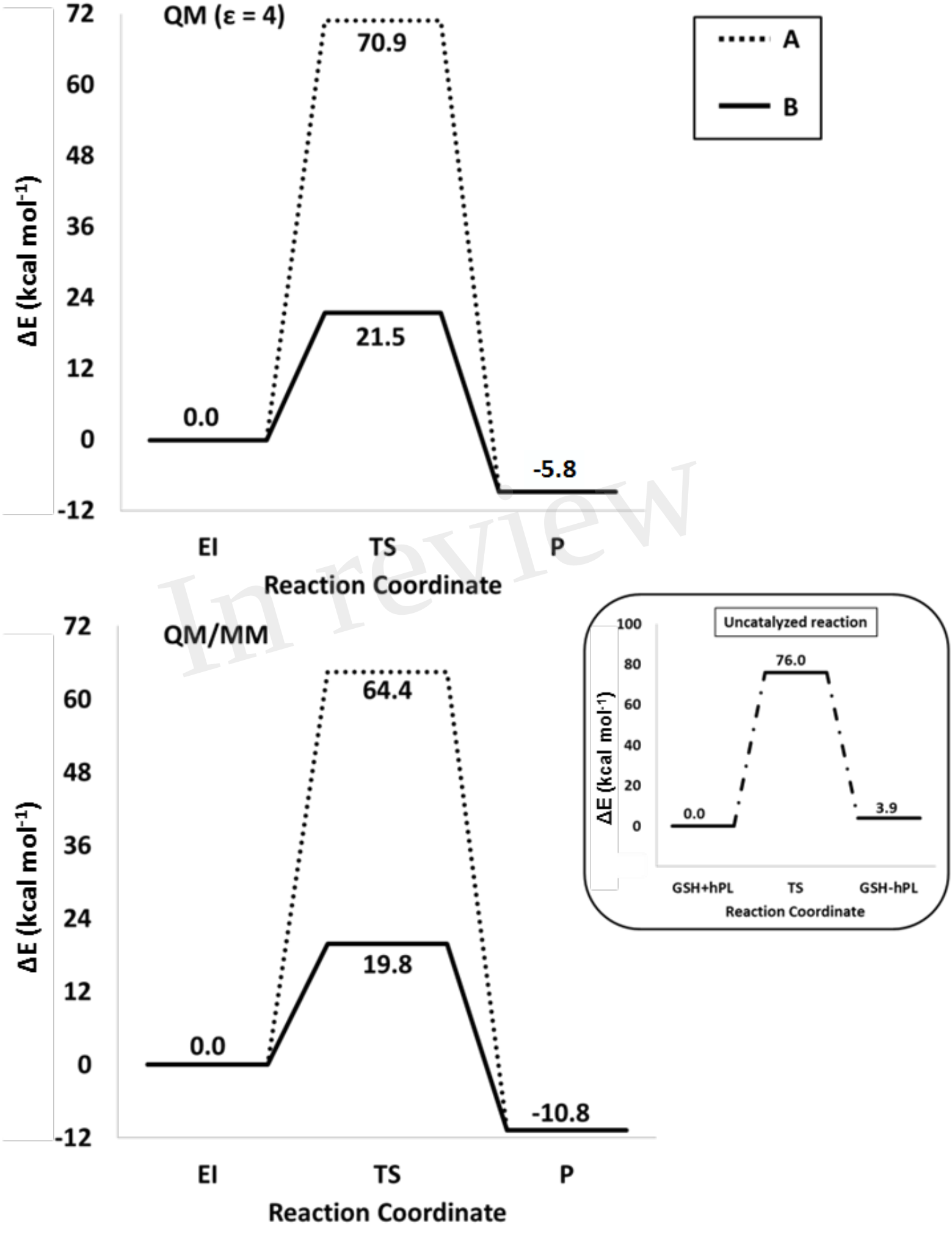


Figure 7.TIF



In review

Figure 8.TIF

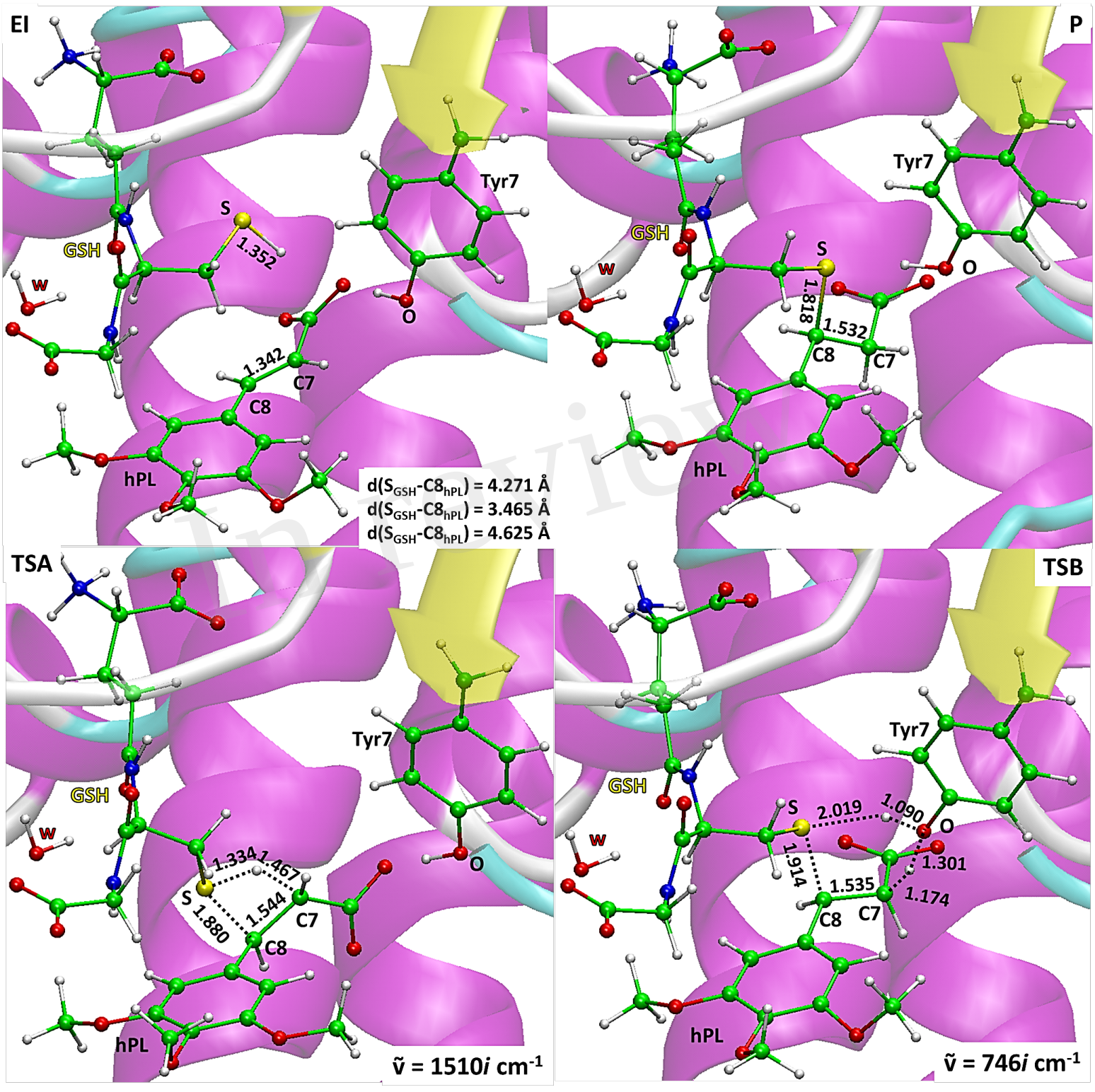

EI
GSH
w
S
1.352
Tyr7
O
1.342
C7
C8
hPL
d(S$_{GSH}$-C8$_{hPL}$) = 4.271 Å
d(S$_{GSH}$-C8$_{hPL}$) = 3.465 Å
d(S$_{GSH}$-C8$_{hPL}$) = 4.625 Å
P
GSH
w
S
1.818
1.532
C8
C7
Tyr7
O
hPL
TSA
GSH
w
1.334
1.467
S
1.880
1.544
C7
C8
Tyr7
O
hPL
ṽ = 1510*i* cm^{-1}
TSB
GSH
w
S
2.019
1.090
O
1.914
1.535
1.301
1.174
C8
C7
Tyr7
hPL
ṽ = 746*i* cm^{-1}

Figure 9.TIF

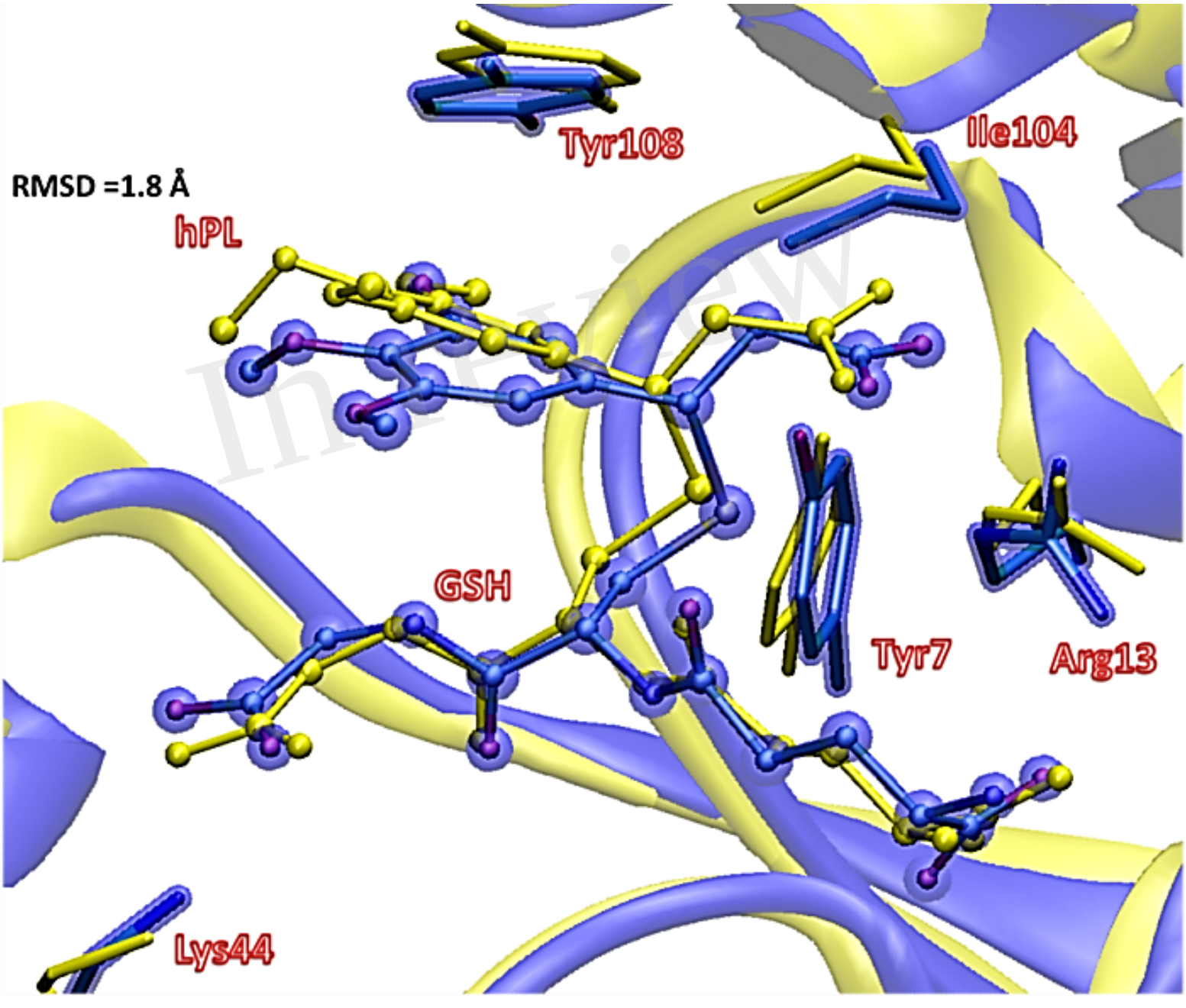
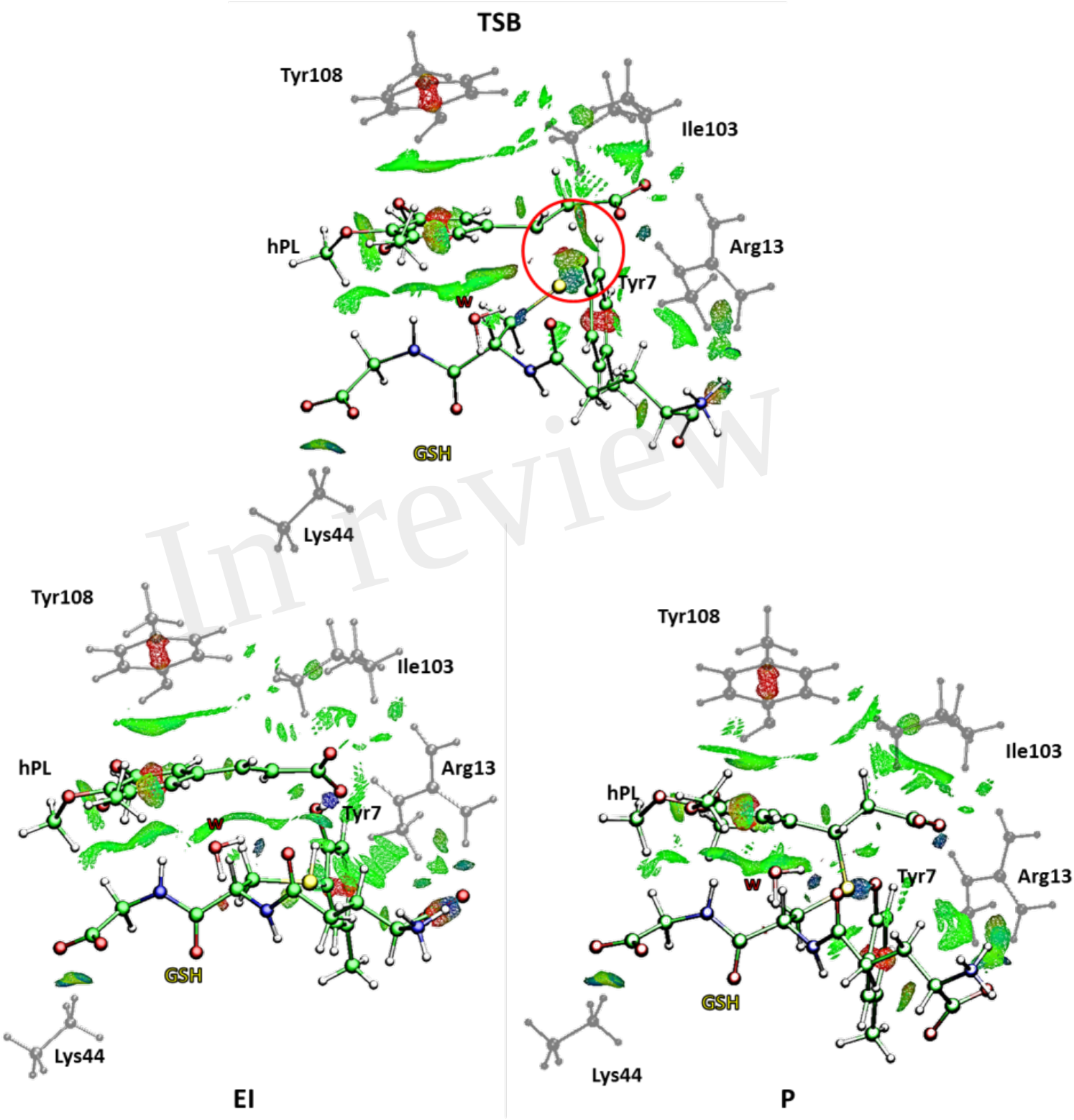


Figure 10.TIF



In review

Wp 1

"How the destabilization of a reaction intermediate affects enzymatic efficiency: the case of human transketolase"

Mario Prejanò, [a] Fabiola E. Medina, [b] Pedro A. Fernandes, [b] Maria J. Ramos, [b] Nino Russo, [a] Tiziana Marino [a]

[a] *Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)*
[b] *UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal*

Manuscript in preparation

How the Destabilization of a Reaction Intermediate Affects Enzymatic Efficiency: the Case of Human Transketolase

Mario Prejanò,[a] Fabiola E. Medina,[b] Pedro A. Fernandes,[b*] Maria J. Ramos,[b] Nino Russo,[a] Tiziana Marino[a*]

[a]*Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)*

[b]*UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal*

Corresponding author

tiziana.marino65@unical.it
pafernan@fc.up.pt

Abstract

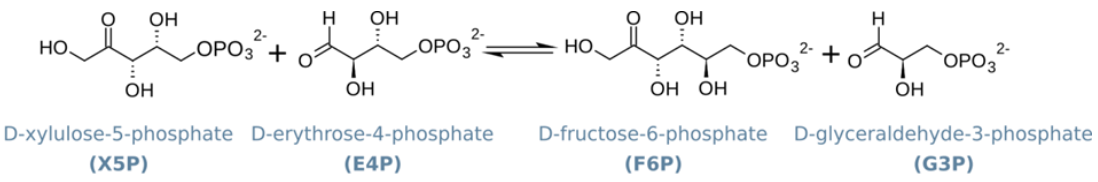
The origin of the catalytic power of enzymes is a subject under debate for decades. It has been proposed that proficient enzymes not only stabilize the rate-determining transition states but also preclude intermediates to become too shallow as to keep the barriers of subsequent steps within an acceptable height. Transketolase (TK) was pointing out as a prototype of the latter. TK is a ThDP-dependent enzyme that catalyzes the transfer of a dihydroxyethyl group from the ketose D-xylulose-5-phosphate (X5P) to the aldose D-erythrose-4-phosphate (E4P), to yield the products D-fructose-6-phosphate (F6P) and D-gliceraldehyde-3-phosphate (G3P). Atomic resolution X-ray rystallography has shown that a crucial intermediate of the catalytic cycle displays a significant, putatively highly-energetic, out-of-plane distortion in a sp^2 carbon of a lytic bond, postulated to lower the barrier of the subsequent step and to contribute for the increase the enzyme's turnover Here we use high-level QM/MM calculations to decipher the TK reaction mechanism, and to show that the origin of the distortion is not steric strain but mostly the establishment of an internal hydrogen bond in the intermediate, whose enthalpy pays-off for the distortion penalty necessarily to its alignment. The distortion associated to hydrogen bonding is present even in small gas-phase models, and in fact its net effect is stabilizing and anti-catalytic. The results help to understand the intrinsic enzymatic machinery behind enzyme's amazing proficiency.

Introduction

Enzymes play an essential role in a broad variety of the biochemical processes. Understanding these processes is an interest, and a major challenge, for the research community.[1, 2] The most popular theory that explains the origin of enzyme's catalytic power was proposed by Pauling in 1948.[3] The underlying idea is that enzymes act by stabilizing more the transition state than the ground state, which is materialized through a higher binding affinity for the former than the latter. Several proposals have been put forward to explain the physical origin of the transition state stabilization, in particular in comparison with the transition state stabilization that takes place in the corresponding uncatalysed aqueous solution reaction.[4-14]

One of these proposals is based in the enzyme promotion of reactive ground state conformations. The proposal suggests that the enzyme rate constant (k_{cat}), and thus the enzyme's efficiency (k_{cat}/K_M), is very dependent on the ground state substrate conformation.[14, 15] A resulting aspect is that a ground state substrate conformation that the binds the enzyme in a conformation that looks like the transition state (which can be seen as "distorted" when compared to the aqueous solution conformation) could hypothetically contribute to the observed kinetic behaviour, lowering the energy to obtain the transition state, and thus increasing k_{cat} and k_{cat}/K_M.[16-18]

Human transketolase (hTK) is a thiamine diphosphate (ThDP)-dependent enzyme that catalyses the two steps reaction shown in **Scheme 1**, consisting in the transfer of a dihydroxyethyl group from the ketose D-xylulose-5-phosphate (X5P) to the aldose D-erythrose-4-phosphate (E4P), to yield the products D-fructose-6-phosphate (F6P) and D-gliceraldehyde-3-phosphate (G3P). The associated catalytic mechanism of such a reaction is proposed to involve a physical distortion in one of its intermediates that lowers the barrier of the subsequent step, contributing to increase the enzyme's k_{cat}. The proposed reaction mechanism suggests that after the formation of the reactive ylide form of ThDP, a nucleophilic attack of the ThDP carbanion (C_2 ¬atom in **Figure 1**) to the carbonyl carbon (C2x)¬ of X5P forms a X5P-ThDP covalent complex. This species is supposed to exhibit a physical distortion (out-of-plane) in relation to the ideal C_2 sp^2 planar geometry, as shown in **Figure 1**.[17,19-21]



Scheme 1. Schematic representation of the general mechanism catalyzed by human transketolase.

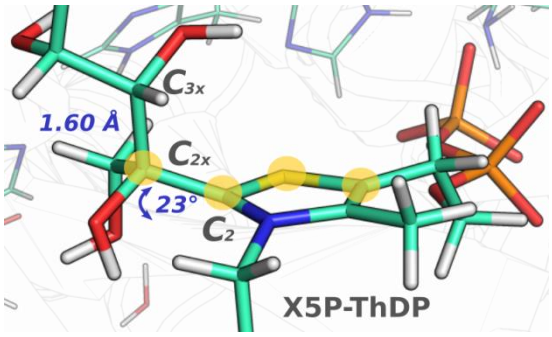


Figure 1. The covalent X5P-ThDP adduct in hTK: Positions highlighted by yellow spheres represent the atoms involved in dihedral angle characterizing the out-of-plane distortion of the X5P-ThDP intermediate. The C_{2x} and the C_{3x} label carbons represent the scissile bond.

A high-resolution (0.97 Å) hTK crystallographic structure recently reported[20] corroborated this hypothesis, raised based in lower-resolution x-ray structures [16], as the C_2 atom of X5P-ThDP has sp^2 hybridization, and thus should present a trigonal planar geometry, but in fact it displayed an out-of-plane distortion of 22°, as shown in **Figure 1**. Furthermore, at this resolution, the X5P-ThDP intermediate exhibits what was considered to be a highly strained, elongated, C_{2x}-C_{3x} scissile bond (1.60 Å), the bond that will be cleaved in the subsequent reaction step.[19] These experimental pieces of evidence point to a catalytic effect, due to ground-state destabilization, or intermediate destabilization. Stimulated by this interesting and unusual enzymatic working mechanism, we have studied the reaction mechanism of hTK employing density functional theory (DFT) both within pure DFT cluster models (QM) and a complete enzyme model described at the hybrid quantum mechanical/molecular mechanical (DFT/MM) level, in order to establish the atomic details of the catalytic cycle and to evaluate the hypothesized catalytic effect originated by the out-of-plane distortion in the X5P-ThDP intermediate (i.e the ground-state destabilization hypothesis). Computer simulations are nowadays the only way to separate the physical components that contribute to the activation free energy and to quantify each of its contributions individually. Additionally, we wanted to understand the possible relationship between the X5P-ThDP distortion and the large bond length of the C_{2x}-C_{3x} scissile bond that is proposed to facilitate the progression over the rate-limiting transition state.

Methods

1. *Enzyme modelling*

The active site molecular models used for QM and QM/MM calculations were obtained starting from the X-ray structure of ThDP covalent bonded with X5P substrate at the resolution of 0.97 Å (PDB id: 4KXV), isolated by Homo Sapiens. [19] hTK is characterized by two homodimers related by a 2-fold rotational symmetry with two identical active sites in the dimer interface. [19-24]

The cofactor is positioned in the catalytic pocket 10 Å deep. Upon binding, X5P engages in several hydrogen bonds with the active site sidechains, as evidenced in **Figure S1**. The substrate phosphate group interacts with Arg100A, Arg318B, Ser345B, His416B and Arg474B. The O4 is involved in hydrogen bonds with Asp424B, His37A and His110A, while O3 is involved in an hydrogen bond with His258A. The O2 interacts with Gln428B while the Glu366 forms an hydrogen bond with the N1 atom of the pyridinium ring; the Lys244A interacts with His258A. One water molecule (W1) makes hydrogen bonds with O1 and with another water molecule (W2), present in the task. These residues have been included during the buildup of both DFT cluster and QMMM models, shown in **Figure 2**.

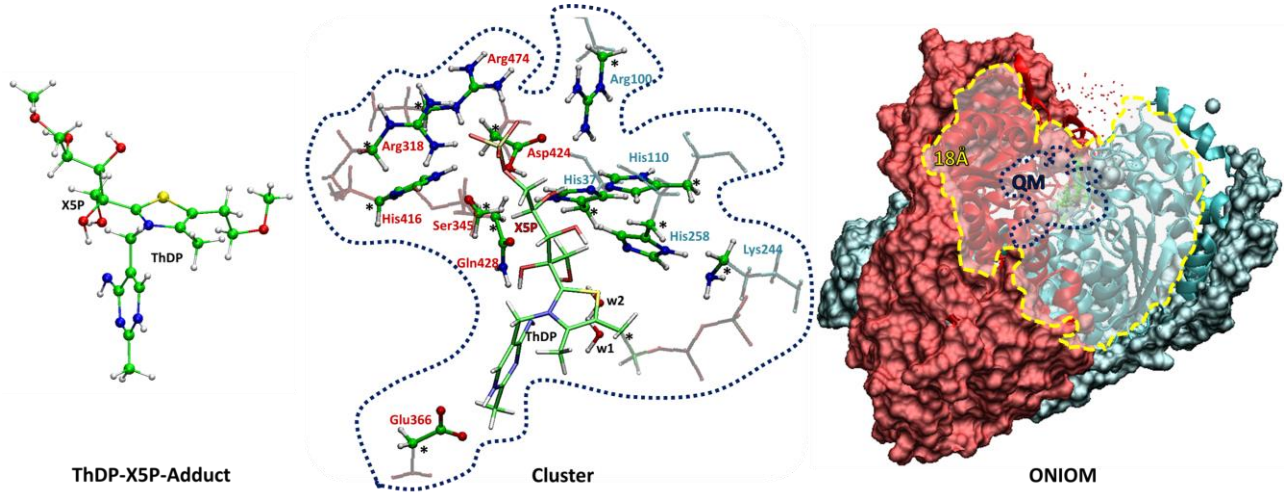


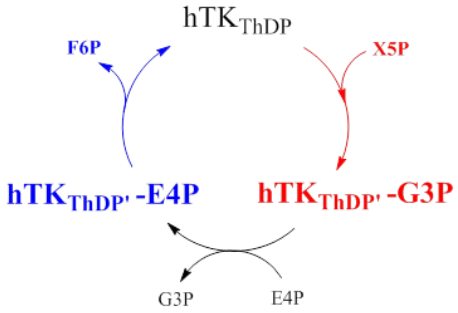
Figure 2. The three models adopted in this work. From left to right: the DFT ThDP-X5P model, the DFT cluster model and the complete enzyme DFT/MM model.

1.1. ThDP-X5P adduct model

The smallest model adopted in the present investigation was obtained by including only the ThDP - X5P covalent intermediate present in the X-ray structure. This model was used to study the physical distortion present in the covalent intermediate without any influence from enzyme scaffold. We have replaced the $–O\text{-}PO_3^{2-}$ and $–CH_2\text{-}O\text{-}P_2O_5^{2-}$, of X5P and ThDP respectively, with $–OCH_3$ and $–CH_3$ groups to facilitate comparison with results coming from a previous work. [16] The final number of atoms was 61. This model was fully described at the DFT level.

1.2. QM Cluster model

The cluster model contains the amino acid residues of the active site described above, the ThDP cofactor and the substrate. The residues and ThDP cofactor were truncated at the α-carbons and the hydrogens of polarizable residues were added considering experimental evidences.[19,23-25] In order to avoid unrealistic rearrangements during the geometry optimizations, the Cα atoms were kept frozen in their positions, (highlighted with an asterisk in **Figure 2**). The residues have been modeled according with the standard cluster model procedure.[26-29] The final models consisted in 184 and 188 atoms (X5P and E4P substrates, respectively), and have a total charge of +1 (see **Scheme 2**).



Scheme 2. Schematic representation of the two reactions occurring in the catalytic process of hTK. The species involved in the first part of the reaction are in red while those in the second one are in blue.

1.3. QM/MM model

The complete protein complexed with ThDP-X5P has been inserted in a rectangular box (124Åx115Åx107Å) of pre-equilibrated TIP3P water molecules[30] and 4 Cl^- counter ions to neutralize the total charge. The protonation states of ionizable residues were predicted by the H++ web server[31] and further compared with the available experimental information (see **Table S1**).[19,23-25] More detailing, mutagenesis and kinetics studies on yeast transketolase suggested the possibility that the dyad His258A-Lys244A, in human transketolase, is involved in acid-base catalysis. According with this proposal, the Lys244A presents an uncommon neutral state.[24] As reported in **Table S1**, for Lys244A was calculated a pka value of 3.989, substantially lower than other values obtained for other residues protonated (pka > 10.000), in agreement with the proposal. Molecular mechanics parameters were derived for the ThDP-X5P adduct. A geometry optimization at the HF/6-31G(d) level was performed and the Restrained Electrostatic Potential[32] (RESP) method was used to derive atomic charges. The Antechamber tools, as implemented in AMBER16 software

package,[33] were used to derive intramolecular parameters and Lennard-Jones parameters, taken from the General Amber Force Field (GAFF)[34]. The file parameters are included in in the **SI**.

The whole model was geometry optimized with standard procedures [35-37] and a progressive heating of 100 ps was performed from 0 to 310. A subsequent MD production phase of 10 ns was carried out, in NPT condition at the temperature of 310 K, monitoring the conformational changes and the variations of relevant geometrical parameters involved in the catalytic reaction mechanism (see **Figure S2** and **Figure S3**). In all the simulations, the SHAKE[38] algorithm, the PME[39] scheme and a cutoff radii of 12 Å were used, in order to adopt integration step of 0.002 ps and evaluate intermolecular interaction.

The QM/MM model was obtained applying the two-layers ONIOM formalism, [40] starting from the minimized structure obtained at ff99SB level of theory,[41] maintaining in the high-level region (QM) the residues selected (His37A, Arg100A, His110A, Lys244A, His258A, Arg318B, Ser345B, Glu366B, His416B, Asp424B, Gln428B and Arg474B) and in the low-level region (MM) the remaining part of the protein, together with the 80 water molecules that were located within 5 Å from the active site cavity. (**Figure 2**) The final model contains 19319 atoms. All atoms within a radius of 18 Å from the QM region were optimized while the water molecules and the remainder of the enzyme were fixed in their positions.

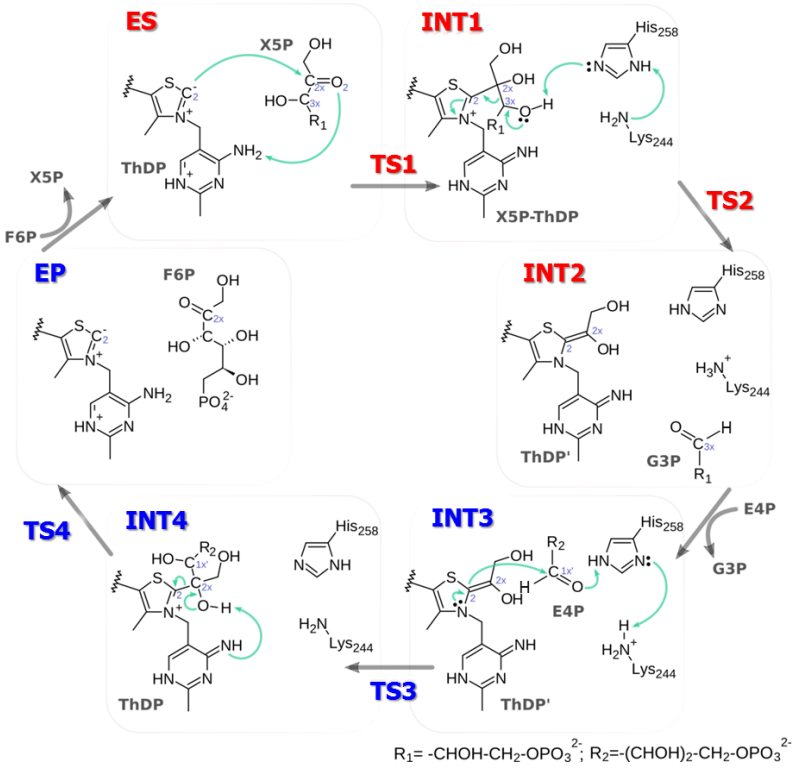
Technical Details

QM and QM/MM calculations were performed using the Gaussian09 software.[42] The active site cluster model was described with the Becke exchange[43] and Lee, Yang and Parr[44] correlation functionals. The 6-31G(d,p) basis set was used for geometry optimizations and the 6-311+G(2d,2p) basis set for single point energy calculations. The ThDP-X5P adduct model was treated both at the MP2[45] level of theory, using the triple-ζ aug-pVTZ basis set, and DFT, applying protocol mentioned above. In the QM/MM calculations, the used QM theoretical level was the same as in the QM cluster model; the ff99SB force field was used in the low-level layer. The Coulomb interactions were evaluated using the electrostatic embedding scheme.[46] The nature of each stationary point found along the Potential Energy Surfaces (PESs) was confirmed by frequency calculations on the optimized structures at the same level of theory. The presence of a single imaginary vibrational frequency confirmed the nature of transition state. Only vibrational temperatures larger than 100 K have been account during the calculation of entropy contribution, according with a validated procedure[47,48a] and successfully proposed in other works.[48] The final energy profiles were obtained adding the Zero Point Energy (ZPE), the empirical D3-dispersion correction [49] and entropy contributions, as reported in **Table S2**. In the cluster model calculations, the protein environment

was simulated via a polarizable continuum model (SCRF-PCM),[50] choosing a dielectric constant of ε=4, as suggested in previous studies.[51-53] Accurate energetic profiles were obtained through a relaxed scan along the dihedral angle involved in the distortion observed in the INT1 (C5-S-C2-C2x). Dihedral angle values from 0° to 23° (with 1º increments) have been monitored and optimized. Finally, the Natural Bon Orbital (NBO)[54] analysis have been performed on all stationary points.

Results and Discussion

On the basis of the available experimental information [19,23-253] we propose the reaction mechanism depicted in the **Scheme 3**. The catalytic cycle includes two steps in which both X5P and E4P substrates are processed.



Scheme 3. Mechanism proposed for the two steps reaction catalyzed by the hTK. The species in red are related to the reactions involving the X5P donor substrate and those in blue to the E4P acceptor substrate.

1. The catalytic mechanism of hTK

1.1. First reaction stage: Catalytic mechanism for the conversion of X5P into G3P

The ES→TS1→INT1 step describes the nucleophilic attack of the C2 carbanion of the reactive ylide form of ThDP on the carbonyl carbon of X5P, producing the X5P-ThDP covalent intermediate

(INT1, **Figure 1**). The leading event in this step was the C2x-C2 bond formation that, starting from the value of 3.263 Å in the QM/MM structure of ES complex (4.832 Å in the QM cluster), reached 1.546 Å (1.600 Å in the QM cluster model) at INT1. In the same step, a proton transfer from the N4' amine group of ThDP to the carbonyl oxygen (O2) of X5P takes place. This step is concerted, but asynchronous with the C2x-C2 bod formation, as TS1 has the C2x-C2 bod partially formed but the N4' proton still resides in the donor. The free energy barrier to be overcome is 7.9 kcal mol^{-1} at QM/MM level (8.4 kcal mol^{-1} for the QM cluster model). **Figure 3** shows the free energy profile for the reaction while in **Figure 4** are depicted the optimized stationary points at QMMM level of theory.

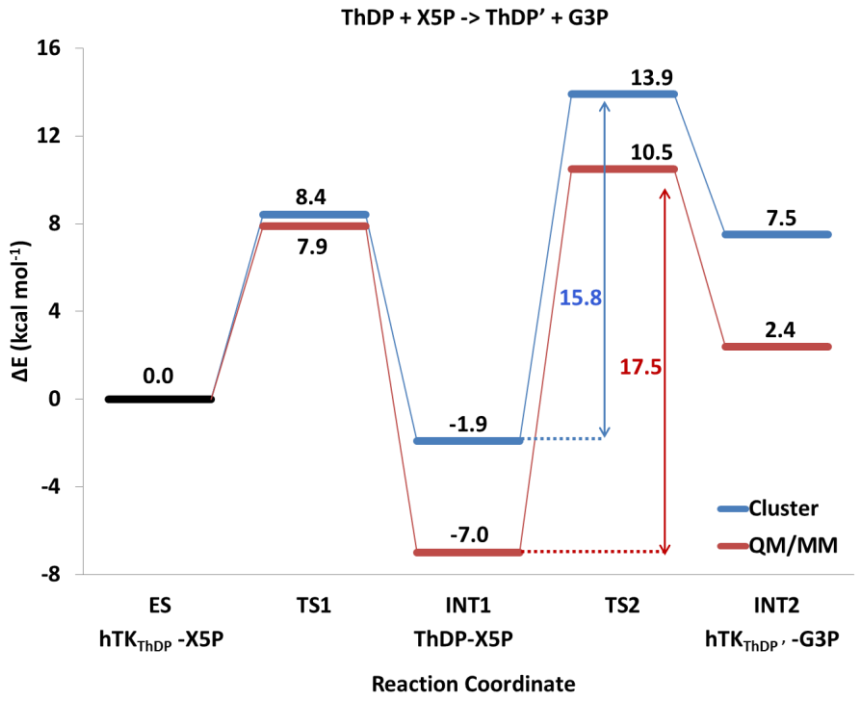


Figure 3. Free energy profile (FEP) for the conversion of X5P into G3P catalyzed by hTK. The red line corresponds to the reaction calculated with a QM cluster model and the blue corresponds to the reaction calculated with a full enzyme model at the QM/MM level.

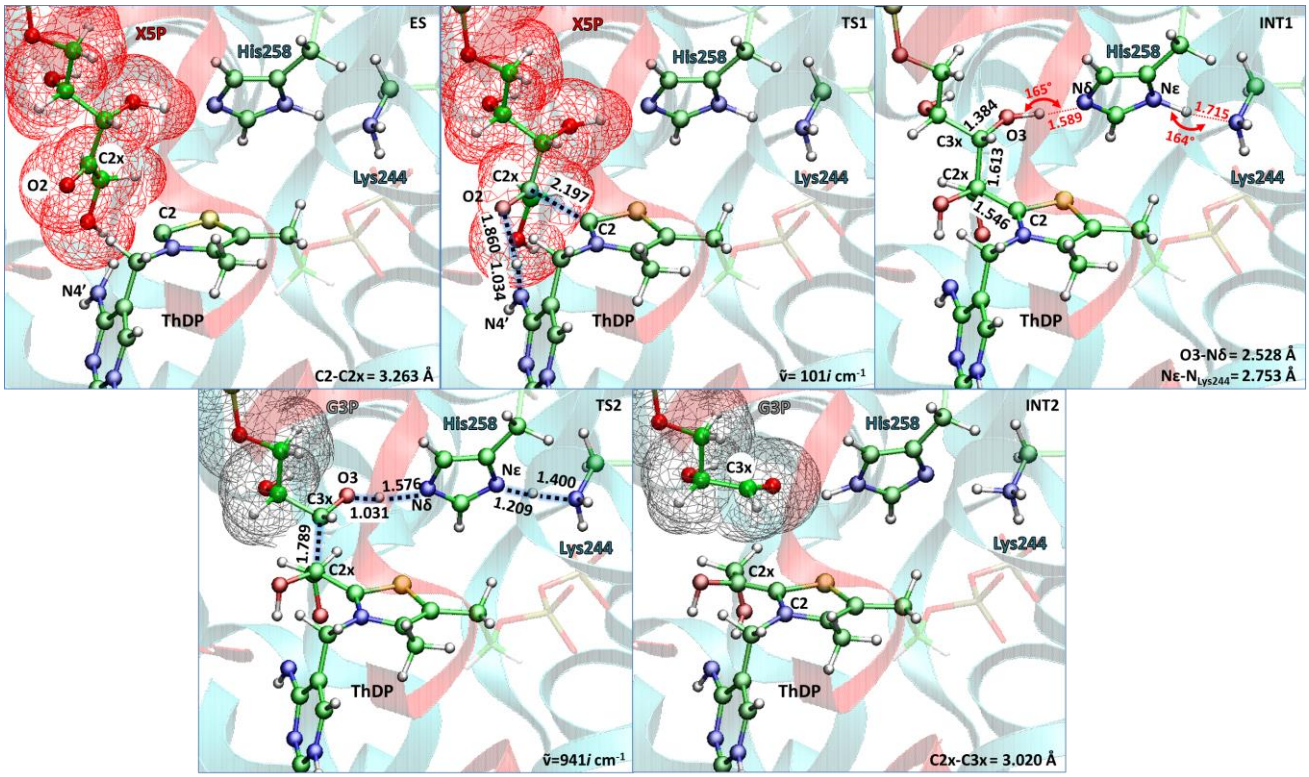


Figure 4. Optimized geometries of the stationary states isolated along the energy surface for the conversion of X5P into G3P calculated at the QM/MM level. Main geometrical parameters are indicated. For TS1 and TS2 the imaginary frequency value is also reported.

Table S3 shows the NBO charges of the most relevant atoms for catalytic process, highlighting that a major depletion of electron charge occurs on the C2 nucleophile atom when moving from ES to TS1. The covalent X5P-ThDP intermediate (INT1) represents the critical point of the reaction since the C2x-C3x scissile bond, with a length of 1.613 Å in the QM/MM model (1.629 Å in the QM cluster model) is significantly longer than typical single C-C bonds (usual bond length around 1.52 Å) , and longer than all other substrate C-C bonds present in the molecular system. In fact, the next step is the cleavage of this bond, to generate the hTKThDP'-G3P (INT2) species, marking the end of the first reaction stage. The latter reaction was simulated, and the barrier for C2x-C3x bond cleavage, seen in **Figure 3**, corresponds to 17.5 kcal mol^{-1} at QM/MM level (15.8 kcal mol^{-1} with the QM cluster model). This step is rate-limiting where the concerted proton transfers take place mediated by His258 asynchronously with the C-C bond breaking. The latter operates as the general base and acid, accepting a proton from O3H group to deliver another one to Lys244 (**Figures 4** and **S5**).

1.2. Second Reaction Stage: Catalytic mechanism of the conversion E4P→F6P.

The second stage of the reaction begins with the E4P acceptor bound to the active site loaded with hTK-ThDP', the latter acting as a donor of a two carbon unit to the former through a two-step processes, whose potential energy profile at QM level is reported in **Figure 5**. In the hTK-ThDP'-E4P complex (INT3), the carbonyl atom (C1x') of E4P lies at 3.616 Å from C2x and in close contact with the His258/Lys244 dyad (O1'-Nδ=2.671Å and Nε-NLys244=2.911 Å), as shown in **Figure S5**. The transition state (TS3) describes the nucleophile addition of C2x to C1x' (1.817 Å), promoted by the lone pair of the thiazolium ring nitrogen atom. This chemical event generates a more negative O1' atom that remains implicated in a series of proton shifts initiated by Lys244 (see **Figure S5**). The atomic charge of O1' decreases by about -0.3e$^-$ in going from the INT3 to TS3 as well as the case of C1x' (**Table S3)** the energy barrier is of only 5.0 kcal mol-1. The product of this reaction (INT4) lies at -15.8 kcal mol^{-1} in relation to INT3 and accounts for the C2x-C1x' bond formation (1.632 Å). This strong exergonicity is consistent with the experimental isolation of the intermediate. [19]

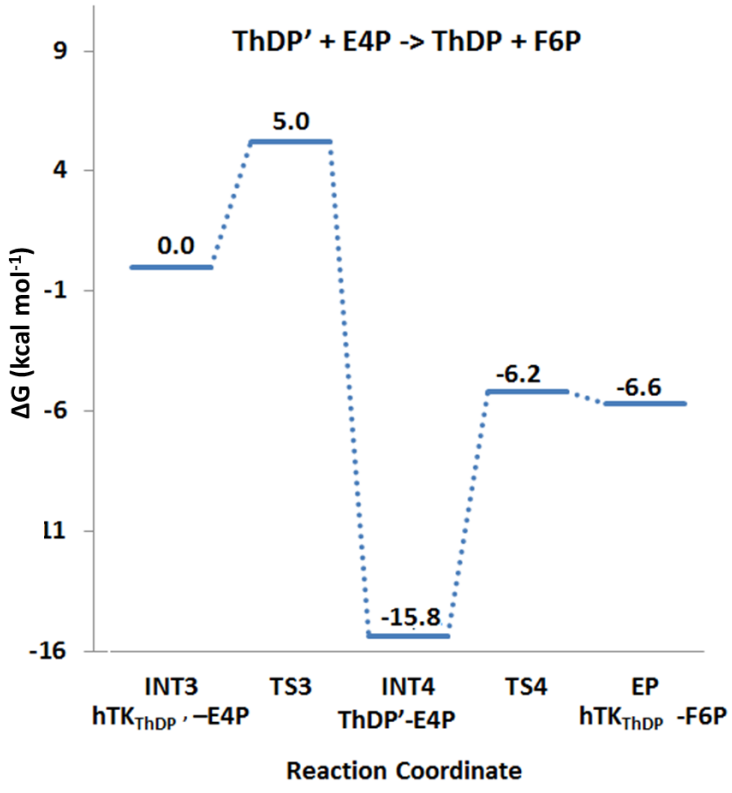


Figure 5. B3LYP/6-311+G(2d,2p) energy profile for the conversion E4P→F6P by hTK.

Interestingly, also in this event a deviation from planarity of the product (about 11°) is observed.

In the subsequent step the breaking of the scissile C2-C2x bond (2.078 Å) occurs and the protonation shift from the O2 to the N4' (O2-H 2.147 Å and H-N4' 1.015 Å) reestablishes the

initial form of the cofactor. This step (TS4) requires about 9.6 kcal mol^{-1}. The final product EP shows the formation of the second sugar having the additional carbon units (F6P) that lies at 4.896 Å from the cofactor. The process is exergonic with an energy of 6.6 kcal mol^{-1}.

Analysis of the X5P-ThDP intermediate (INT1).

The out-of-plane distortion of the C2-C2x bond connecting ThDP to X5P at INT1 was proposed to enhance the reactivity of INT1, reducing the barrier for the subsequent step[19], as detailed in the Introduction section above. To analyze and quantify this effect, postulated to represent an important catalytic effect and a paradigm of catalysis through "intermediate destabilization", we built up an intentionally reduced additional molecular model, consisting only in the X5P-ThDP intermediate, stripped of the enzyme scaffold, to decouple the effect of the physical distortion from any other energy contribution emanating from the enzymatic system. (see **Figure 6**).

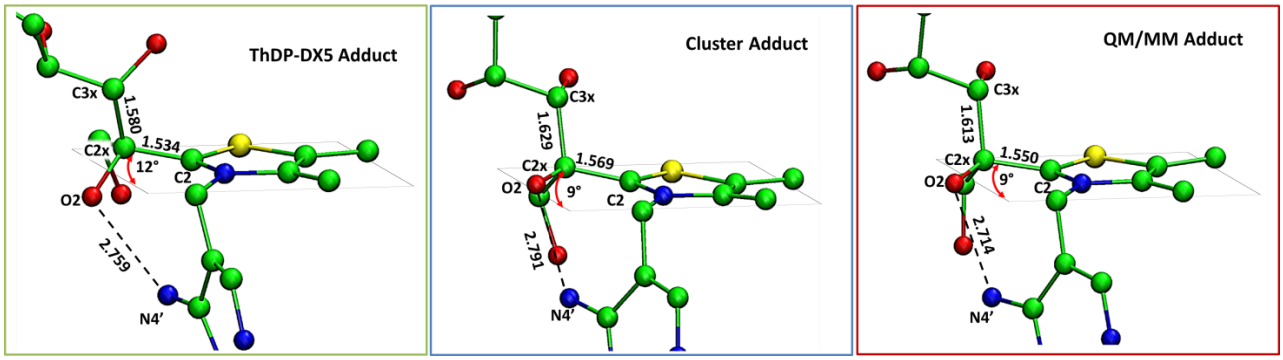


Figure 6. The out-of-plane distortion of the C2-C2X bond in the three used models. All models show a similar 9º-12º distortion. The ThDP-X5P model is free to rotate around the C2x-C3x bond and avoid the repulsion between the O2 and N4' atoms without clashing with the protein scaffold, eliminating the out-of-plane distortion, but the distorted geometry is still more favorable because it allows for the establishment of an internal O2-N4' hydrogen bond.

Upon free geometry optimization, all computational models displayed a quantitatively very similar distortion (**Figure 6**), and in qualitative (but not so much quantitative) agreement with the X-ray structure.

A deviation from the planarity of 9° is obtained by taking into account the effects of the neighboring residues, partially as in the QM cluster or fully in the ONIOM. The minimal ThDP-DX5 model have shown a distortion of 12º. The distortion in the X-ray structure is geometrically similar, but larger (22 º). In addition, the QM and QM/MM models have both the cofactor thiazolium and aminopyrimidine rings in a V-type arrangement, with an interatomic distance

between O2 and N4' of 2.791 and 2.714 Å. This well matches with the strained (X-ray) value of 2.790 Å.19 The O2-N4' distance has been pointed out as the responsible for the out-of-plane distortion, as a planar structure would place the O2 and N4' in a too close, repulsive contact19.

The minimal ThDP-X5P model, stripped from the whole enzyme provides similar results, with an O2-N4' distance of 2.759 Å.

It may look surprising that the ThDP-X5P structure, that is free to rotate around the C2x-C3x bond without clashing with the protein, avoiding in this way a bad contact between O2 and N4' still shows an out-of-plane distortion . The fact that it does indicates that the distortion is overall energetically favorable, as it allows the fine-tuning of the distance and angle of the O2-N4' internal hydrogen bond. Altogether, the establishment of a more favorable hydrogen bond seems to compensate for the distortion penalty. As such, the energy penalty due to the distortion has necessarily to be small.

We have performed relaxed PES scan as function of the C5-S-C2-C2x dihedral angle. This dihedral has the value of 0º for a non-distorted planar structure, and a value of 22º in the X-ray distorted intermediate. We monitored the C2-C2x, C2x-C3x and O2-N4' distances as a function of the dihedral angle. The results are shown in **Figure 7**.

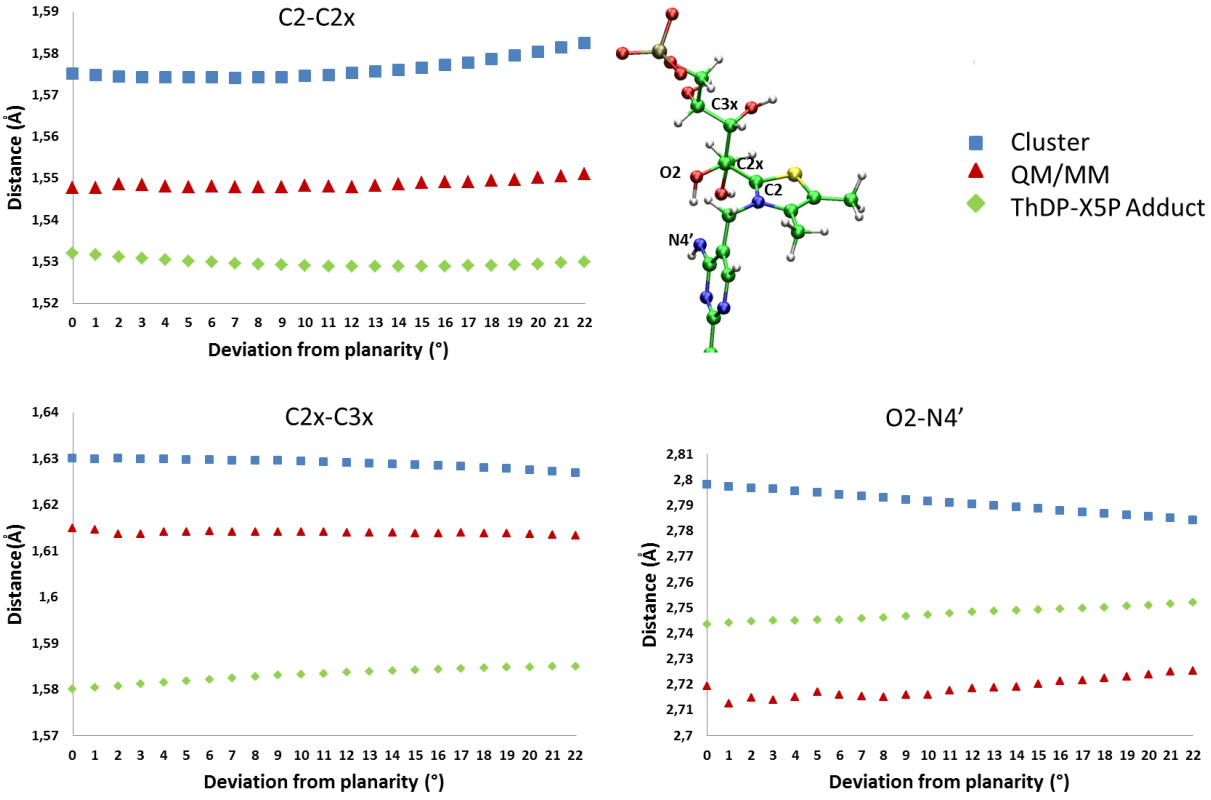


Figure 7. The C2-C2x, C2x-C3x and O2-N4' distances as function of the C5-S-C2-C2x dihedral, from 0° to 22°

The distances obtained by the relaxed scans are in excellent agreement with the experimental x-ray structure.[19] More importantly, there are no substantial changes in the three examined geometrical parameters, as the distortion grows from 0º to 22º. Maybe the most important observation is that the C2x-C3x bond is streched (in relation to a typical CC single bond) across all the spectrum of out-of-plane distortions. Therefore, it becomes clear the physical distortion of the intermediate is not the cause of the uncommonly elongated C2x-C3x scissile bond, contrarily to what has previously believed.

In fact, a similar length covalent bond has been recently reported occurring in an intermediate for the reaction mechanism of enoyl reductase.[22, 23] The energy difference between the planar and the distorted geometries is very small (1.5 kcal mol^{-1} in the full enzyme model, 0.4 kcal mol^{-1} in the two smaller models, **Figure 8a**), as it is the difference in energy between the computational and the experimental distortion (less than 2.6 kcal mol^{-1} when moving from the computational distortion to the experimental distortion). The flatness of the PES may be the reason why the experimental and computational distortions show differences of 11º-14º; small differences in the medium (crystal vs. solution), temperature, and /or small inaccuracies in the computations may easily translate in small energy differences, but significant geometric differences through the flat PES. The influence of the O2-N4' hydrogen bond on the out-of-plane distortion has been evaluated, through the replacement of N4' group with hydrogen atom in the ThDP-X5P model, generating the ThDP-$X5P_{no_N4'}$ model (see **Figure 8b**). The structure optimized to a perfect planar configuration, losing the earlier physical distortion, definitively showing that, in the computational model, the favorable establishment of the internal O2-N4' hydrogen bond is at the origin of the physical distortion.

The ThDP-$X5P_{no_N4'}$ molecular model allows to decouple the energy penalty coming from the physical distortion from the stabilization energy coming from the internal O2-N4' hydrogen bond. The physical distortion energy penalty corresponds to the energy difference between the structure of the ThDP-$DX5_{no_N4'}$ model with the dihedral angle at 0º and at 12º (or between 0º and 22º if we want to use the experimental value for the distortion), and it corresponds to 1.7 and to 2.0 kcal mol^{-1}. To calculate the stabilizing effect of the internal hydrogen bond we first calculate the difference between the energy of the ThDP-DX5 model at angle at 0º and at 12º (or between 0º and 22º if we want to use the experimental value for the distortion), which encompasses both the contributions of the physical distortion and hydrogen bonding, and then we subtract the physical distortion penalties calculated with the ThDP-$DX5_{no_N4'}$ model. In this way, we obtain a value of 0.7 and 1.0 kcal/mol for the stabilizing effect of the internal O2-N4' hydrogen bond.

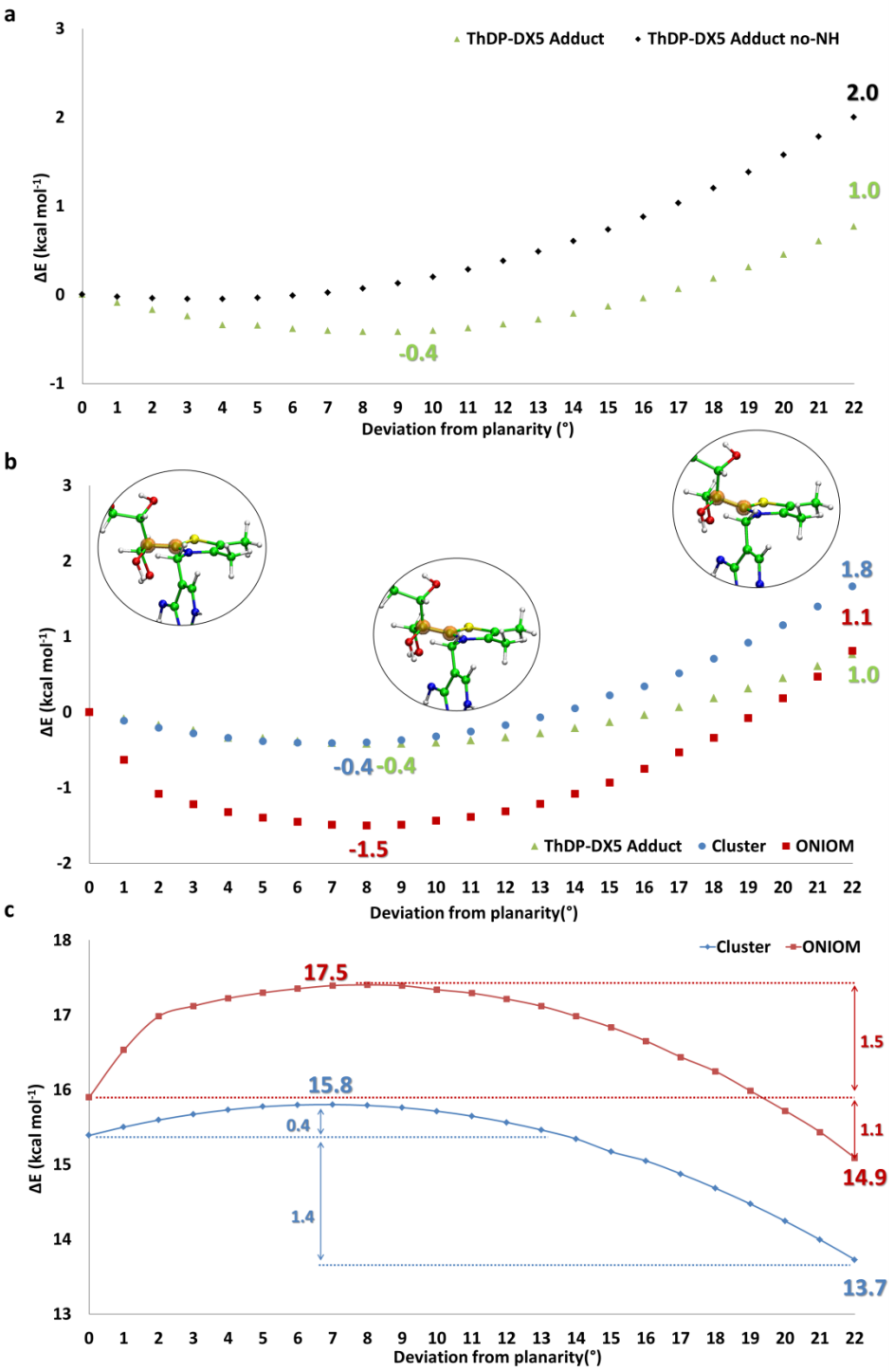


Figure 8. (a). Potential energy profile as function of out-of-plane distortion of the C2-C2x bond, from 0° to 22°; b) Potential energy profile as a function of the out-of-plane distortion of the C2-C2x bond in the adduct ThDP-X5P model (green) and in the same model but with the N4' group replaced by hydrogen (black); c). Limiting step energy barriers (INT1-TS2) as function of deviation from planarity from 0° to 22°.

Considering how energy of INT1 variates as function of the distortion, we note that barrier at QM/MM level decreases from 17.5 kcal mol-1 (for deviation of 9 degrees) to 14.9 kcal mol-1 for

deviation experimentally observed (22°), [19] which accounts for an increase of the kcat of about two orders of magnitude. (**Figure 8c**)

A previous study on a similar reduced ThDP-X5P system, at the DFT (B3LYP/6-31G(d)) level, indicated that forcing the planarity of the ThDP-X5P complex without re-organizing its structure, and without forming the intramolecular 2-OH-N4' hydrogen bond (due to the specific orientation of the hydrogen atoms that was specifically modeled) might involve a penalty of ~20 kcal mol^{-1} [16]. A full relaxation of the same model led to the establishment of an internal hydrogen bond between the 4-OH group of the substrate and the negative phosphate, making this more realistic, relaxed model, to lie in a different, much shallower, local minimum than the strained structure where this hydrogen bond was absent, a fact that precluded the evaluation of the real distortion energy 16. The fully relaxed model shown an out-of-plane deviation of 10º. Finally, in the same study, a X5P-thiazolium model was geometry optimized, and surprisingly displayed an out-of-plane distortion of 9º 16, despite the model lacking the N4' amino group. This result is difficult to explain , and it is in stark contrast with our model where the N4' amine was replaced by an hydrogen atom, which became perfectly planar upon geometry optimization.

In summary, our calculations suggest that the physical distortion seen in the ThDP-X5P covalent intermediate do not raise its energy, as they allow for the establishment of an internal hydrogen bond that pays off the distortion penalty. As such, this distortion does not represent a true catalytic effect. The distortion is not the cause of the stretching of the scissile C2-C2x bond, as this bond is longer than usual CC single bonds irrespectively of the physical distortion applied to the ThDP-X5P covalent intermediate.

Conclusions

This work was devoted to understand of the catalytic mechanism of hTK at the atomic level, and to investigate if a physical distortion of a key intermediate induced by the enzyme scaffold raised its energy to a point that it represented a relevant catalytic effect by lowering the next, rate-limiting barrier. The mechanism was concluded to take place in four steps: in the first and second steps the ThDP attack the substrate (X5P) favoring the release of G3P product, while in the third and fourth steps, the second substrate (E4P) interacts with the ThDP'in order to form the second product (F6P).The physical distortion in the ThDP-X5P covalent intermediate was shown not to result from the limited capacity of relaxation of the adduct due to its trapping in the enzyme scaffold, as models stripped from the enzyme, and free to relax, still show the distortion. Instead, we suggest that the distortion is induced by a favorable intramolecular hydrogen bond, whose stabilization energy pays off the price for distorting the planarity of the intermediate structure. We also suggest that the

abnormal long length of the scissile C2-C2X bond is not a consequence of the physical distortion, as previously believes, as planar, relaxed structures still exhibit a very elongated bond. It seems that the physical origin of the enzymatic efficiency of hTK is still a well-kept secret. Further studies are needed to understand the origin of its catalytic power.

References

1. Hammes, G. G.; Benkovic, S. J.; Hammes-Schiffer, S. *Biochemistry* **2011,** *50* (48), 10422.
2. Amyes, T. L.; Richard, J. P. *Biochemistry* **2013,** *52* (12), 2020.
3. Pauling, L. *Nature* **1948,** *161* (4097), 707.
4. Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1971,** *68* (8), 1678.
5. Menger, F. M.; Glass, L. E. *J. Am. Chem. Soc.* **1980,** *102* (16), 5404.
6. Bruice, T. C. *Acc. Chem. Res.* **2002,** *35* (3), 139.
7. Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. USA.* **1985,** *82* (8), 2225.
8. Cleland, W. W.; Kreevoy, M. M. *Science* **1994,** *264* (5167), 1887.
9. Agarwal, P. K. *J. Am. Chem. Soc.* **2005,** *127* (43), 15248.
10. Zhang, X. Y.; Houk, K. N. *Acc. Chem. Res.* **2005,** *38* (5), 379.
11. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. *Chem. Rev.* **2006,** *106* (8), 3210.
12. Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004,** *303* (5655), 186.
13. Benkovic, S. J.; Hammes-Schiffer, S.. *Science* **2003,** *301* (5637), 1196.
14. Giraldo, J.; Roche, D.; Rovira, X.; Serra, J. *FEBS Lett.* **2006,** *580* (9), 2170.
15. Zhang, X.; Houk, K. N. *Acc. Chem. Res.* **2005,** *38* (5), 379.
16. Asztalos, P.; Parthier, C.; Golbik, R.; Kleinschmidt, M.; Hubner, G.; Weiss, M. S.; Friedemann, R.; Wille, G.; Tittmann, K. *Biochemistry* **2007,** *46* (43), 12037.
17. Asztalos, P.; Parthier, C.; Golbik, R.; Kleinschmidt, M.; Hubner, G.; Weiss, M. S.; Friedemann, R.; Wille, G.; Tittmann, K. *Biochemistry* **2007,** *46* (43), 12037.
18. Lehwess-Litzmann, A.; Neumann, P.; Parthier, C.; Lüdtke, S.; Golbik, R.; Ficner, R.; Tittmann, K. *Nat. Chem. Biol.* **2011,** *7*, 678.
19. White, J. K.; Handa, S.; Vankayala, S. L.; Merkler, D. J.; Woodcock, H. L. *J. Phys. Chem. B* **2016,** *120* (37), 9922.
20. Ludtke, S.; Neumann, P.; Erixon, K. M.; Leeper, F.; Kluger, R.; Ficner, R.; Tittmann, K. *Nat. Chem.* **2013,** *5* (9), 762.

21. Meshalkina, L. E.; Solovjeva, O. N.; Kochetov, G. A. *Biochemistry (Mosc.)* **2011,** *76* (9), 1061.
22. Rosenthal, R. G.; Ebert, M.-O.; Kiefer, P.; Peter, D. M.; Vorholt, J. A.; Erb, T. J. *Nat. Chem. Biol.* **2014,** *10* (1), 50.
23. Rosenthal, R. G.; Vogeli, B.; Quade, N.; Capitani, G.; Kiefer, P.; Vorholt, J. A.; Ebert, M. O.; Erb, T. J. *Nature Chemical Biology* **2015,** *11* (6), 398.
24. Tittmann. K. *Bioorgan. Chem.* **2014**, *57*, 263.
25. Mitschke, L.; Parthier, C.; Schröeder-Tittmann, K.; Lüdtke, S.; Tittmann, K. *J. Biol. Chem.* **2010**, 285 (41),31599.
26. Siegbahn, P. E. M.; Himo, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*(3), 323-336.
27. Himo, F. *J. Am. Chem. Soc.* **2017,** *139*(20), 6780.
28. Piazzetta, P.; Marino, T.; Russo, N.; Salahub, D. R. *ACS Catal.* **2015**, *5*(9), 5397.
29. Prejanò, M.; Marino, T.; Russo, N. *Chem. Eur. J.* **2017**, 23, 8652.
30. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *79*, 926.
31. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. *Nucleic acids research* **2012**, *40*(W1), W537.
32. Bayly, C. I.; Cieplak, P.; Cornell, W.D., Kollman, P.A. *J. Phys. Chem.* **1993**, *97*(40), 10269.
33. Case, D. A.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.;Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.;LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. (2017), AMBER 2017, University of California, San Francisco.
34. Wang, J.; Wolf, R.M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comp. Chem.* **2004**, *25*(9), 1157.
35. Prejanò, M,; Marino, T.; Russo, N. *Frontiers in chemistry* **2018**, *6*, 249.
36. Medina, F. E.; Neves, R. P. P.; Ramos, M. J.; Fernandes, P. A. *Phys Chem Chem Phys* **2017**, *19*, 347.
37. Sousa, S. F. M.; Ribeiro, A. J.; Neves, R. R. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; M. J. Ramos, M. J. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1281.

38. Ryckaert, J-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 1977, *23*(3), 327-341.
39. Ewald, P. P. *Ann Phys-Berlin* **1921**, *64*(3), 253.
40. Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*(50), 19357.
41. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
42. Frisch, M. J.; G. W., Trucks, Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2013
43. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
44. Lee, C. T.; Yang, W T.; Parr, R. G. *Phys Rev B* **1988**, *37*(2), 785.
45. Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Let.t* **1990,** *166*(3), 275.
46. Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A. Jr.; Morokuma, K; Frisch, M. J. *J. Chem. Theory Comput.* **2006,** *2*(3), 815.
47. Grimme, S. *Chem. Eur. J.* **2012**, *18*, 9955.
48. (a) Neves, R. P. P.; Fernandes, P. A.; Ramos, M. J. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4724. (b) Medina, F. E.; Neves, R. P. P.; Ramos, M. J.; Fernandes, P. A. *ACS Catal.* **2018**, *8*, 10267.
49. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456.
50. Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.
51. Warshel, A., In Computer Modeling of Chemical Reactions in Enzymes and Solutions; Wiley, New York, 1991.

52. Ramos, M. J.; Fernandes, P. A. *Acc Chem Res* **2008**, *41*(6), 689.

53. Piazzetta, P.; Marino, T.; Russo, N. *Inorg. Chem.* **2014**, *53*(7), 3488.

54. E. D. Glendening, A. E. Reed, J. E. Carpenter,F.Weinhold, NBO, version 3.1.

How the Destabilization of a Reaction Intermediate Affects Enzymatic Efficiency: the Case of Human Transketolase

Mario Prejanò,[a] Fabiola E. Medina,[b] Pedro A. Fernandes,[b*] Maria J. Ramos,[b] Nino Russo,[a] Tiziana Marino[a*]

[a]*Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Arcavacata di Rende (CS) (Italy)*

[b]*UCIBIO, REQUIMTE, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal*

Corresponding author

tiziana.marino65@unical.it

pafernan@fc.up.pt

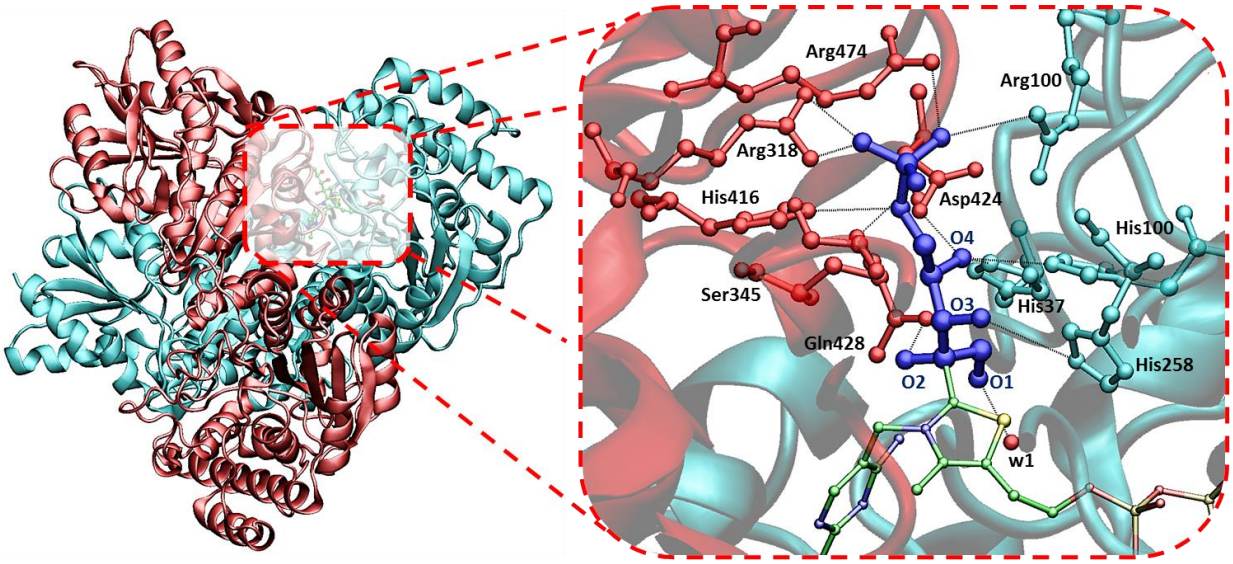


Figure S1. Cartoon representation of hTK (left) . Residues interacting with the X5P substrate (dark blue). Amino acids are colored in cyan and red, belonging to the A chain and B chain, respectively.

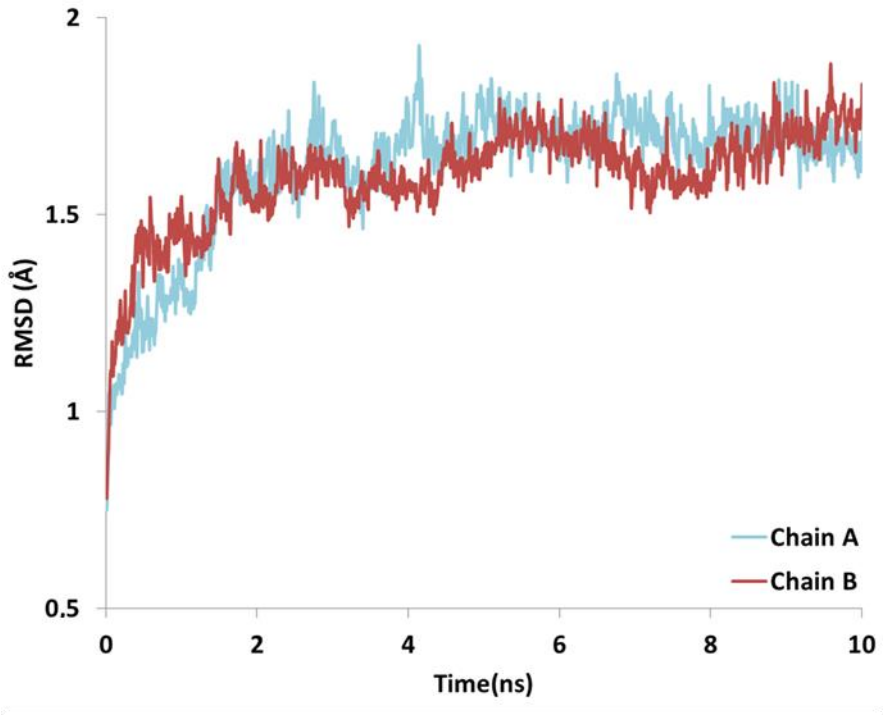


Figure S2. The Root Mean Square Deviation (RMSD) profile of the hTK's two chains.

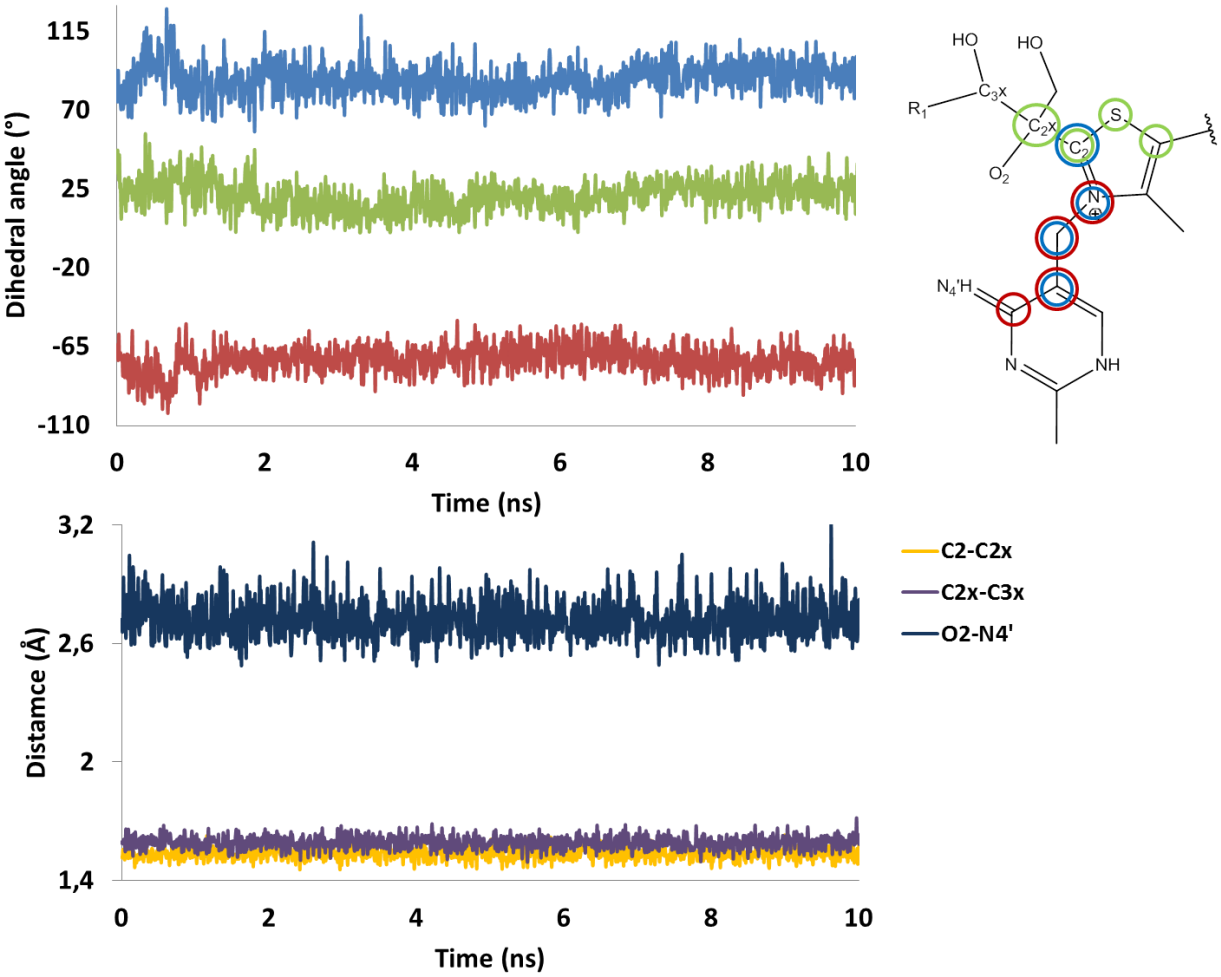


Figure S3. Dihedral variations and important distances, long 10 ns of classical Molecular Dynamics.

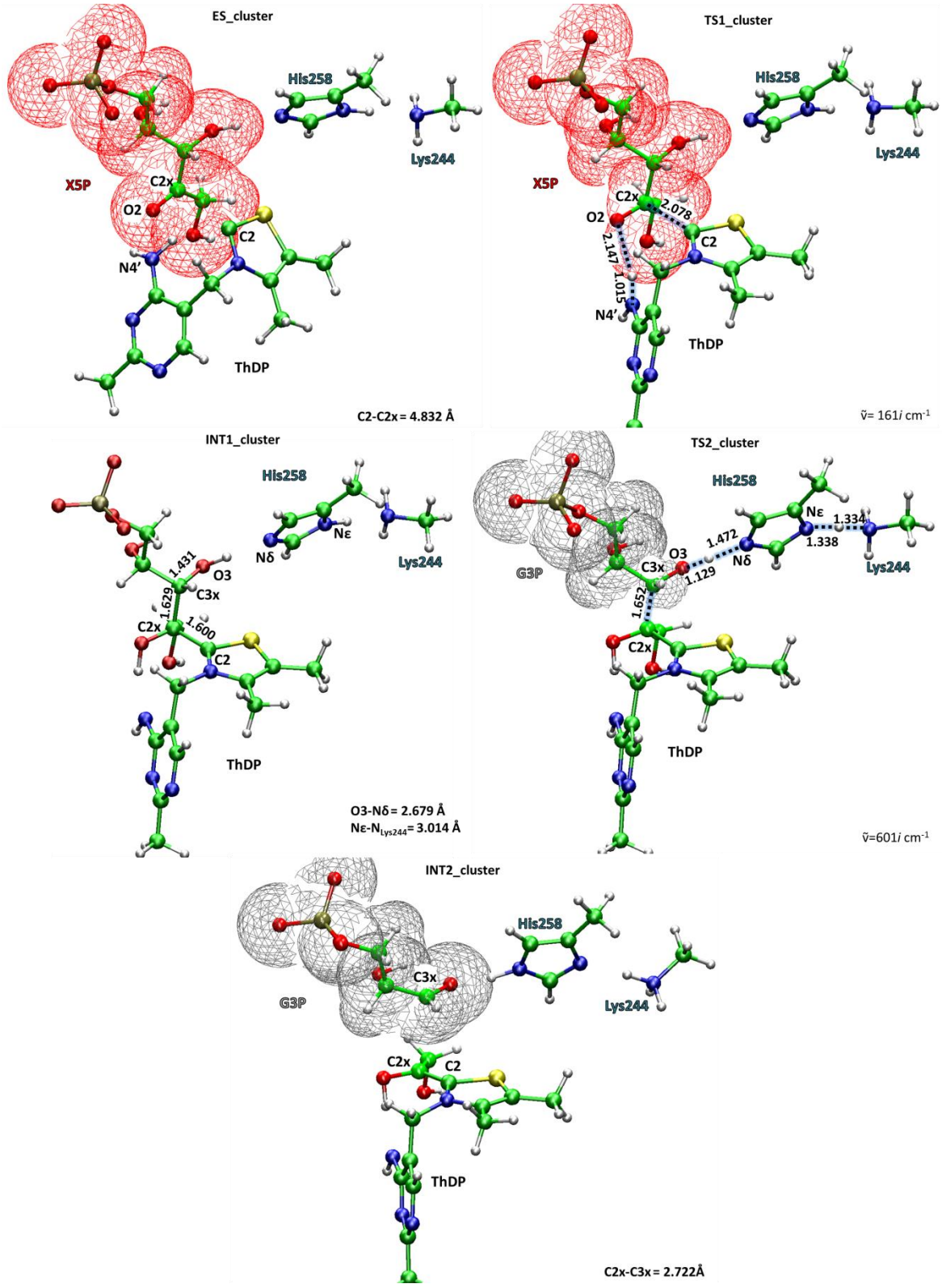


Figure S4. B3LYP/6-31G(d,p) optimized geometries in QM Cluster for first path of reaction.

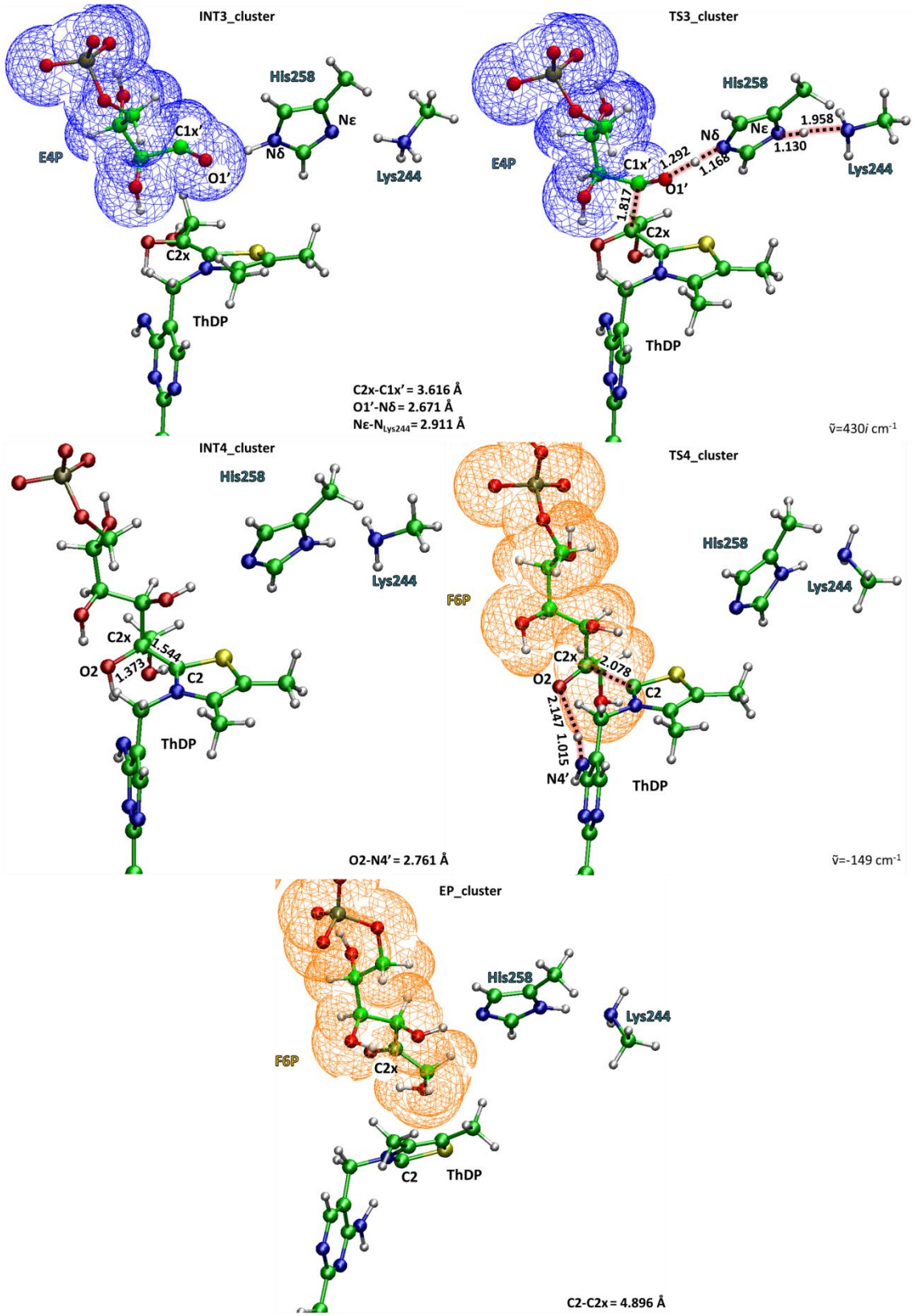


Figure S5. B3LYP/6-31G(d,p) optimized geometries in QM Cluster for second path of reaction.

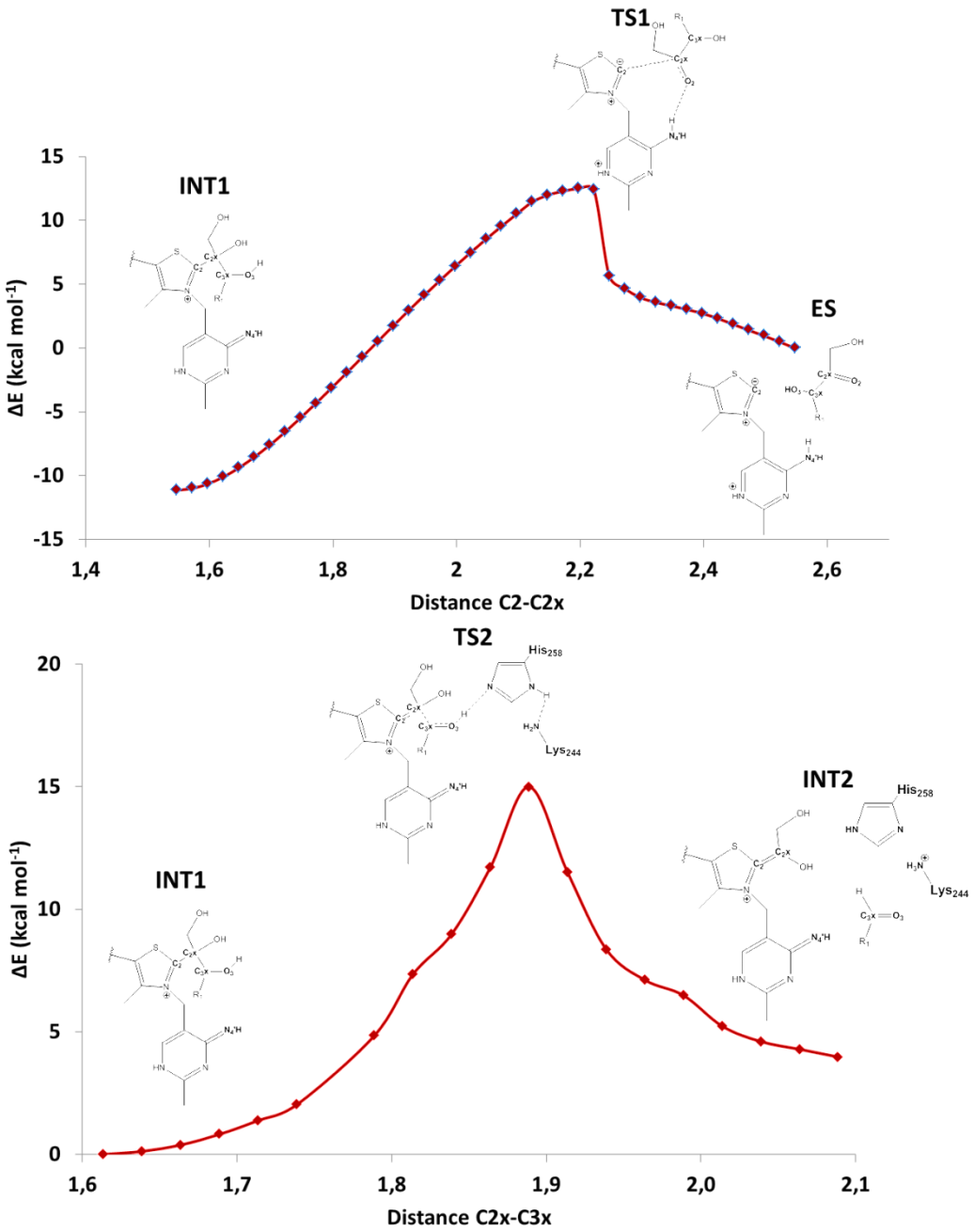


Figure S6. Potential Energy Surfaces, calculated at B3LYP/6-31G(d,p):ff99SB level of theory, in order to detect transition states in ONIOM calculations.

Table S1. Calculated pKa for ionizable residues of hTK. Amino acids fully protonated (positively charged) or deprotonated (negatively charged) are highlighted in blue and red, respectively. In bold are represented the groups maintained in QM portion.

Residue	pKa	Residue	pKa	Residue	pKa	Residue	pKa
TYR-4	>12.000	TYR-173	>12.000	LYS-327	>12.000	LYS-499	10.336
HIS-5	6.015	LYS-174	10.689	HIS-330	8.293	ASP-500	3.226
LYS-6	10.97	ASP-176	0.969	ASP-333	2.754	ASP-501	<0.000
ASP-8	1.944	ASP-183	<0.000	ARG-334	>12.000	HIS-513	7.023
LYS-11	9.886	ARG-186	>12.000	ASP-339	<0.000	GLU-514	<0.000
LYS-16	10.293	ASP-191	3.209	ASP-341	4.308	GLU-520	3.849
ASP-17	0.229	HIS-197	5.179	LYS-343	>12.000	LYS-523	10.764
ARG-21	>12.000	ASP-200	3.698	GLU-349	3.158	LYS-524	10.581
ARG-23	>12.000	TYR-202	>12.000	LYS-352	9.784	GLU-525	1.816
HID-37	<0.000	LYS-204	10.072	LYS-353	11.124	LYS-526	10.777
CYS-41	>12.000	ARG-205	>12.000	GLU-354	2.149	ARG-530	>12.000
CYS-42	>12.000	CYS-206	>12.000	HIS-355	4.372	ASP-533	<0.000
GLU-46	<0.000	GLU-207	4.249	ASP-357	2.687	LYS-538	>12.000
HIS-54	<0.000	HIS-212	3.755	ARG-358	>12.000	ASP-541	1.472
ARG-57	>12.000	ASP-217	1.926	GLU-361	1.692	ARG-542	>12.000
TYR-58	>12.000	HID-219	<0.000	CYS-362	>12.000	LYS-543	>12.000
LYS-59	11.664	GLU-222	3.587	TYR-363	>12.000	ASP-547	2.61
ASP-62	1.837	GLU-223	3.26	GLU-366	4.854	ARG-550	>12.000
ARG-64	>12.000	CYS-225	>12.000	CYS-376	>12.000	LYS-553	10.847
HIS-67	<0.000	LYS-226	10.557	ARG-379	>12.000	ARG-555	>12.000
ASP-69	<0.000	LYS-232	10.932	ARG-381	>12.000	GLU-560	<0.000
ARG-70	>12.000	HIS-233	5.946	CYS-386	>12.000	ASP-561	<0.000
LYS-75	>12.000	LYS-241	11.434	ARG-395	>12.000	HID-562	<0.000
HIS-77	<0.000	LYS-244	3.989	ASP-398	<0.000	TYR-563	>12.000
TYR-83	>12.000	ARG-246	>12.000	ARG-401	>12.000	TYR-564	>12.000
GLU-88	<0.000	GLU-252	0.619	GLU-407	4.557	GLU-565	4.777
GLU-94	4.344	ASP-253	<0.000	CYS-413	>12.000	GLU-570	2.562
GLU-96	2.691	LYS-254	11.274	HIS-416	<0.000	GLU-579	2.89
ARG-100	11.014	GLU-255	3.722	CYS-417	>12.000	HIS-586	10.084
LYS-102	11.102	HIS-258	<0.000	GLU-423	4.917	ARG-591	>12.000
ASP-106	1.886	LYS-260	10.31	ASP-424	10.17	ARG-594	>12.000
ASP-108	0.313	LYS-264	9.785	GLU-432	5.576	LYS-597	10.837
HID-110	2.588	GLU-268	4.244	ASP-433	<0.000	GLU-600	3.17
LYS-114	>12.000	GLU-273	3.795	ARG-438	>12.000	LYS-603	11.053
ASP-119	<0.000	TYR-275	9.987	TYR-447	>12.000	ASP-608	3.924
CYS-133	>12.000	LYS-281	>12.000	ASP-450	<0.000	ARG-609	>12.000
TYR-137	>12.000	LYS-282	10.212	GLU-455	<0.000	ASP-610	3.639
LYS-140	>12.000	LYS-283	>12.000	LYS-456	>12.000	ARG-617	>12.000
TYR-141	>12.000	GLU-291	2.675	GLU-459	2.998		
ASP-143	<0.000	ASP-292	4.808	LYS-465	10.896		
LYS-144	11.549	ASP-297	4.127	CYS-468	>12.000		
TYR-147	>12.000	ARG-302	>12.000	ARG-471	>12.000		
ARG-148	>12.000	TYR-309	>12.000	ARG-474	>12.000		
TYR-150	>12.000	LYS-310	>12.000	GLU-476	1.999		
CYS-151	>12.000	ASP-313	2.32	TYR-481	>12.000		
ASP-155	9.013	LYS-314	9.516	GLU-485	2.994		
GLU-157	<0.000	ARG-318	>12.000	ASP-486	5.074		
GLU-160	4.783	LYS-319	9.679	LYS-493	>12.000		
GLU-165	4.327	TYR-321	>12.000	LYS-497	10.04		

Table S2. Energy contributions extrapolated for each stationary points. All values are in kcal mol^{-1}. In parenthesis are reported values calculated in for QMMM model.

	$\Delta E_{B3LYP-D3}$	Δ ZPE	$-T\Delta S$	ΔG
ES	0.0	0.0	0.0	0.0
TS1	6.1 (9.4)	1.6 (0.4)	0.7 (-1.9)	8.4 (7.9)
INT1	-4.9 (-7.1)	3.4 (1.6)	-0.4 (-1.5)	-1.9 (-7.0)
TS2	13.6 (15.3)	0.6 (-3.5)	-0.3 (-1.3)	13.9 (10.5)
INT2	7.7 (2.5)	0.7 (1.4)	-0.9 (-1.5)	7.5 (2.4)

	$\Delta E_{B3LYP-D3}$	Δ ZPE	$-T\Delta S$	ΔG
INT3	0.0	0.0	0.0	0.0
TS3	2.9	2.2	-0.1	5.0
INT4	-18.4	3.0	-0.4	-15.8
TS4	-7.8	2.6	-1.0	-6.2
EP	-8.4	2.8	-1.0	-6.6

Table S3. Variation of charges, in NBO population analysis, of relevant atoms considered during the mechanism. In red are highlighted the values obtained for ONIOM model while the remaining are for QM Cluster calculations.

	C2		C2x		O2		C3x		O3	
ES	-0.25	-0.26	0.57	0.57	-0.58	-0.57	-0.03	0.00	-0.86	-0.83
TS1	-0.06	-0.13	0.45	0.49	-0.78	-0.71	0.01	0.03	-0.86	-0.83
INT1	0.13	0.12	0.22	0.21	-0.80	-0.79	0.05	0.09	-0.87	-0.83
TS2	0.13	0.12	0.21	0.21	-0.81	-0.79	0.06	0.13	-0.89	-0.82
INT2	-0.03	0.00	0.21	0.21	-0.74	-0.78	0.29	0.33	-0.74	-0.68

	C1x'	O1'
INT3	-0.05	0.21	-0.99	0.43	-0.60
TS3	0.16	0.16	-0.93	0.13	-0.88
INT4	0.16	0.19	-0.93	0.05	-0.81
TS4	-0.1	0.46	-0.80	-0.01	-0.77
EP	-0.25	0.58	-0.62	-0.03	-0.84

Table S4. Energy contributions extrapolated by PESs for each stationary point. In red are highlighted the values obtained for ONIOM model while the remaining are for QM Cluster calculations.

Species	E_{B3LYP} (a.u.)	E_{ZPE} (a.u.)	E_{D3} (a.u.)	E_{total} (a.u)
ES	-5015.682088	1.536765	-0.340509	-5014.145323
	-5066.062893	77.194208	-0.222180	-4990.738904
TS1	-5015.672332	1.539387	-0.355454	-5014.132945
	-5066.044235	77.194826	-0.230321	-4990.723260
INT1	-5015.689869	1.542214	-0.356432	-5014.147655
	-5066.068333	77.196703	-0.228614	-4990.747778
TS2	-5015.660466	1.537797	-0.351806	-5014.122529
	-5066.042222	77.188604	-0.228851	-4990.720104
INT2	-5015.669797	1.537937	-0.359427	-5014.132000
	-5066.050986	77.196389	-0.227163	-4990.732741
INT3	-5130.246438	1.56832	-0.359996	-5128.678118
TS3	-5130.24175	1.57187	-0.367991	-5128.66988
INT4	-5130.275683	1.573099	-0.369348	-5128.702584
TS4	-5130.258854	1.572388	-0.369411	-5128.686466
EP	-5130.259893	1.572705	-0.35701	-5128.687188

Wp 2

“Mechanistic insights of hydrolytic activity into a de novo functional protein framework.”

Mario Prejanò, Isabella Romeo, Tiziana Marino, Nino Russo

Manuscript in preparation

Mechanistic insights of hydrolytic activity into a *de novo* Functional protein framework.

Mario Prejano', Isabella Romeo, Tiziana Marino*, Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy.

Corresponding author

tiziana.marino65@unical.it

Abstract

The reaction mechanism of an artificial esterase obtained by engineering the functional catalytic triad (Cys-His-Glu) into a fully *de novo* designed coiled-coil homo-heptameric peptide assembly (CC-Hept) is proposed on the basis of a combined molecular dynamics (MD) and hybrid quantum mechanics/molecular mechanics (QM/MM) investigation. The preliminary MD simulations have been performed on both unbound and bound to the model substrate (p-nitrophenyl acetate) protein, thus supporting the stability of *de novo* protein architecture. Two reaction pathways have been deeply analyzed at QM/MM level evidencing the rate determining step in agreement with the observed kinetics evidences. The role played by the water molecules and the catalytic triad has been highlighted at atomistic level. Our results should be useful for future developments of more selective and efficient engineered enzymes.

Introduction

The hydrolytic enzymes are required both in pathological and physiological fields as well in industrial applications including food, paper, and ethanol production.[1] Nevertheless, most of the naturally occurring enzymes are expensive, characterized by wide specificities, activity range under restrictive temperature and pH conditions, and in some cases also associated with a toxic profile.[2] Consequently, it is highly recommended the development of inexpensive, eco-friendly, and best performing enzymes. For this reason, protein scientists led the way in protein design and engineering.[3] In this regard, the ability to build from scratch proteins able to synthesize small molecules could result in a positive impact in different areas such as green chemistry, medicine, industrial biotechnology and drug production.[4]

Despite the huge number of protein sequences that can be generated with modern genomics sequencing approaches, the experimental efforts are too far from exploring them all. At this purpose, the design and the engineering of *de novo* protein by using computational techniques is an important guide in providing solutions to new protein-structure/function targets. Recently, several successes in proteins design have been reached for the generation of non-ribosomial peptide synthetates,[5] and new members of the di-iron carboxylate proteins from the Due Ferri family,[6] and cofactor-binding proteins.[7]

Moreover, the structure-based design of hyperstable coiled coils and α-helical barrels considering a different number of helix bundles was one of the most investigated fields.[8] The peculiarity of these structures resides in the willingness to install accurately functional residues in specific environments due to its barrel topology with layers of hydrophobic residues which define the hydrophobic core.[9] Furthermore, these systems can accommodate several mutations in their channels, allowing polar and charged side chains to be included and specific substrates to be recognized.[10] However, considering all the possible combinations of the functional esterase active sites, it's not easy to identify the correct position of the nucleophile and oxyanion in the same catalyst machinery also due to its instability. Probably, the strictness of the protein backbone proposes the use of peptidic amides for oxyanion hole stabilization in *de novo* protein, thus ensuring the preorganization of the active site.[11] The preorganization of the active site to a structure which is able to stabilize the transition state represents one of the pivotal roles in the

design of this kind of proteins. To make this, it is, therefore, appropriate to guarantee the needed rigidity for the catalytic residues with respect to the plasticity of the whole protein. According to that, the initial "lock and key" paradigm shifts in an "induced fit" hypothesis in order to ensure a wider range of motion for the enzyme, thus promoting the catalysis.[12]Indeed, the protein dynamics could be considered the cardinal rule to act on the rate enhancement of enzymes, due to thermal conformational fluctuations which cause variations in the electrostatic environment between the active site and the remaining enzyme.[13,14]

Herein, we started from an attractive study on unnatural "Frankenstein" protein marked by the esterase activity obtained through multiple mutations (Cys-His-Glu) introduced into a fully engineered coiled-coil homo-heptameric peptide (CC-Hept). [15,16]

The hydrolytic activity was achieved without the aid of metal cofactors and the hard-won stability in *de novo* architectures allowed the accommodation of organic substrates in the resulted channel 8 Å in diameter.[15]

Nevertheless, the catalytic efficiency was lower than a recent *in silico* designed metalloprotein,[17] thus underlining the need to improve both efficiency and selectivity of CC-Hept. To achieve this goal, understanding in detail the catalytic mechanism of active Cys-His-Glu triad into Frankenstein protein becomes crucial.

To the best of our knowledge, insights on the mechanism for hydrolytic activity into *de novo* protein, unlike the natural Cys/Ser hydrolases,[18]are not yet explained in the literature.

Therefore, in this work, we combine molecular dynamic simulations (MDs) and the hybrid quantum mechanics/molecular mechanics (QM/MM) calculations to rationalize the hydrolytic mechanism of the model substrate, such as p-nitrophenyl acetate (*p*NPA), catalyzed by a fully *de novo* biocatalysts useful for further enzymes designs to be applied in biotechnology and biochemical synthesis.

Methods

Structure preparation and MD simulations. Starting from the X-ray model of the heptameric coiled coil apo-form, deposited in the Protein Data Bank (PDB) with the code 5EZC,[15] we performed the first step of MDs by using AMBER16[19] package in order to investigate the

structural stability. We built the initial configuration by using xLEaP[20] and Amber ff99SB[21] as force field. The protonation states for all amino acids at physiological pH were calculated according to H++ server.[22] The system was solvated in an orthorhombic box with a buffer of 10 Å, using TIP3P water molecules as the solvent model, meanwhile the counter ions (Cl^-) were added within 2 Å of the protein in order to neutralize the net charge. The solvated structure was first minimized by applying positional harmonic restraints on all atoms (50 kcal mol^{-1} $Å^2$) using 5000 steps of steepest descent[23] followed by 5000 steps of conjugate gradient (CG). After minimization, the unrestrained system was heated up to 310K for 100 ps using the Langevin thermostat. 100 ns of MDs were performed under the following conditions: integration step of 2 fs coupling SHAKE algorithm, NPT ensemble at 1 bar pressure using the Berendsen barostat[24] with a time constant $\tau p = 2.0$ ps. The Particle mesh Ewald summation method[25] was employed in order to find the electrostatic potential and the long-range electrostatic interactions were calculated with 12 Å cut-off distance. To analyse all simulations, we monitored the time evolution of the mass-weighted Root Mean Square Deviations (RMSD), calculated on the backbone atoms of the protein using *cpptraj* from AmberTools 16[26] and we computed the Root Mean Square Fluctuations (RMSF) over all the residues. In order to select different conformations of the system, we performed RMSD-based clustering of the whole trajectory according to the relaxed complex scheme (RCS) docking protocol. [27,28] After removing overall rotation and translations by RMS-fitting the Cα atoms' positions of the trajectory, we applied the average linkage clustering algorithm, implemented in *cpptraj*, identifying 10 significant conformations. These structural insights allowed both the local and global flexibility of the protein and decreased the computational cost for the RCS docking procedure.[29]

Molecular docking calculations. Molecular docking experiments of the *p*NPA to *de novo* proteins were carried out by using AutoDock Vina[30]and 10 output poses were generated. Box centroid was determined by a geometric center of the seven Cys involved into the substrate-binding region and a box of 32 Å size for X, Y and Z was used for grid point generation. For each representative structure, the best docking model with the lowest binding energy was selected. In order to choose the best pose for further investigation geometric criteria were

adopted as follows: a distance cut-off equal to 5 Å for S_{Cys} - C_{pNPA} and S_{Cys} - $N_{\varepsilon His}$, and 3 Å for $N_{\delta His}$ - O_{Glu}.

*p*NPA MM parameters were calculated from single molecule optimization at HF/6-31G(d) level of theory. The General Amber Force Field (GAFF)[31] and the Restrained Electrostatic Potential (RESP)[32] methods were employed to obtain intramolecular, Lennard-Jones parameters and atomic charges, respectively (Table S1). The second step of MDs was applied to the *p*NPA-bound protein with the same protocol above mentioned. During the whole trajectory, we monitored the occupancy of non-bonding interactions between *p*NPA and the protein. After clustering the trajectory, we selected the best complex as input file in QMMM calculations using the geometrical filter above described.

QMMM ONIOM. The QMMM model input file was prepared considering entire structure, the *p*NPA and water molecules within 5 Å from protein, to apply two layers ONIOM formalism.[33] The QM portion of 141 atoms (*p*NPA, $Cys18_{A-G}$, $His22_{A-G}$, $Glu25_{G}$ along with three water molecules close to active site, see Figure 1) had a total charge equal to -1; the remaining atoms (7056) were treated at MM level of theory. During the optimizations, water molecules (1261) around the protein were kept frozen in their positions, applying the procedure for single conformation PES studies. [34,35]

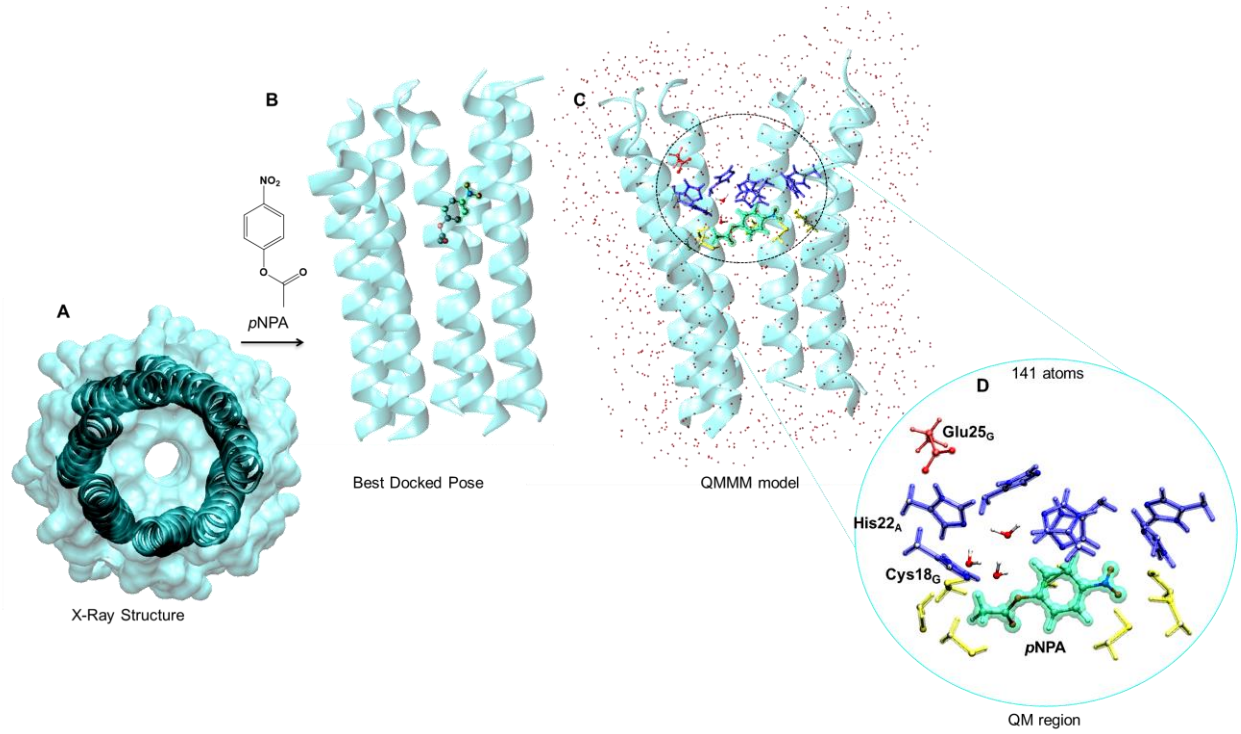


Figure 1. (**A**) Crystal structure of *de novo* designed heptameric coiled coil CC-Hept-Cys_{18}-His_{22}-Glu_{25} selected for preliminary 100 ns of the apo-form protein MDs. (**B**) 2D Representation of unnatural substrate *p*NPA adopted for the docking approach and best docked pose into the active site of *de novo* protein. *p*NPA into the binding cavity is shown as green ball and stick and the protein is represented as cyan cartoon. (**C**) Extrapolated QM/MM model from 100 ns of MDs. Water molecules within 5 Å around the protein are shown as red dots; Cys, His and Glu residues are distinguished with colors of yellow, blue and red ball and stick, respectively. (**D**) Focus on QM region residues involved in the proposed catalytic mechanisms with explicitly retained water molecules for a total of 141 atoms.

ONIOM calculations were carried out using Gaussian 09[36] software package, applying the electrostatic embedding scheme in order to evaluate Coulomb interactions between MM and QM layers.[37] The link atom approach from Gaussian 09 to saturate the valences that resulted from the truncation of bonds across the DFT and MM layers has been applied. The B3LYP/6-31+G(d,p):FF99SB level of theory have been used for all geometry optimizations. The B3LYP[38,39] functional coupled with the before mentioned basis set has proven to be adequate for purpose.[40] We performed relaxed geometry optimization calculations for all stationary points of

the potential energy surfaces. In order to confirm their nature (no imaginary frequencies for minima and one imaginary frequency for transition states (TS) in the QM layer) and to obtain the thermochemical analysis, frequencies calculation was performed at the same level of theory. The larger 6-311+G(2d,2p) basis set has been adopted to obtain more accurate electronic energies of all the stationary points. The final energies are Gibbs energies including zero point energy (ZPE) and dispersion corrections.[41] The natural bonding orbitals' (NBO)[42] as implemented analysis has been carried out on each stationary point intercepted along the reaction coordinate.

Results

With the aim of investigating the behavior of CC-Hept-Cys-His-Glu, we performed 100 ns of MDs of the apo-form protein. After system equilibration, the trend of the RMSD calculated on the backbone of the enzyme during the simulation showed that the protein remained stable for the whole trajectory, fluctuating around an average RMSD value of 2.924 Å (Figure 2C).

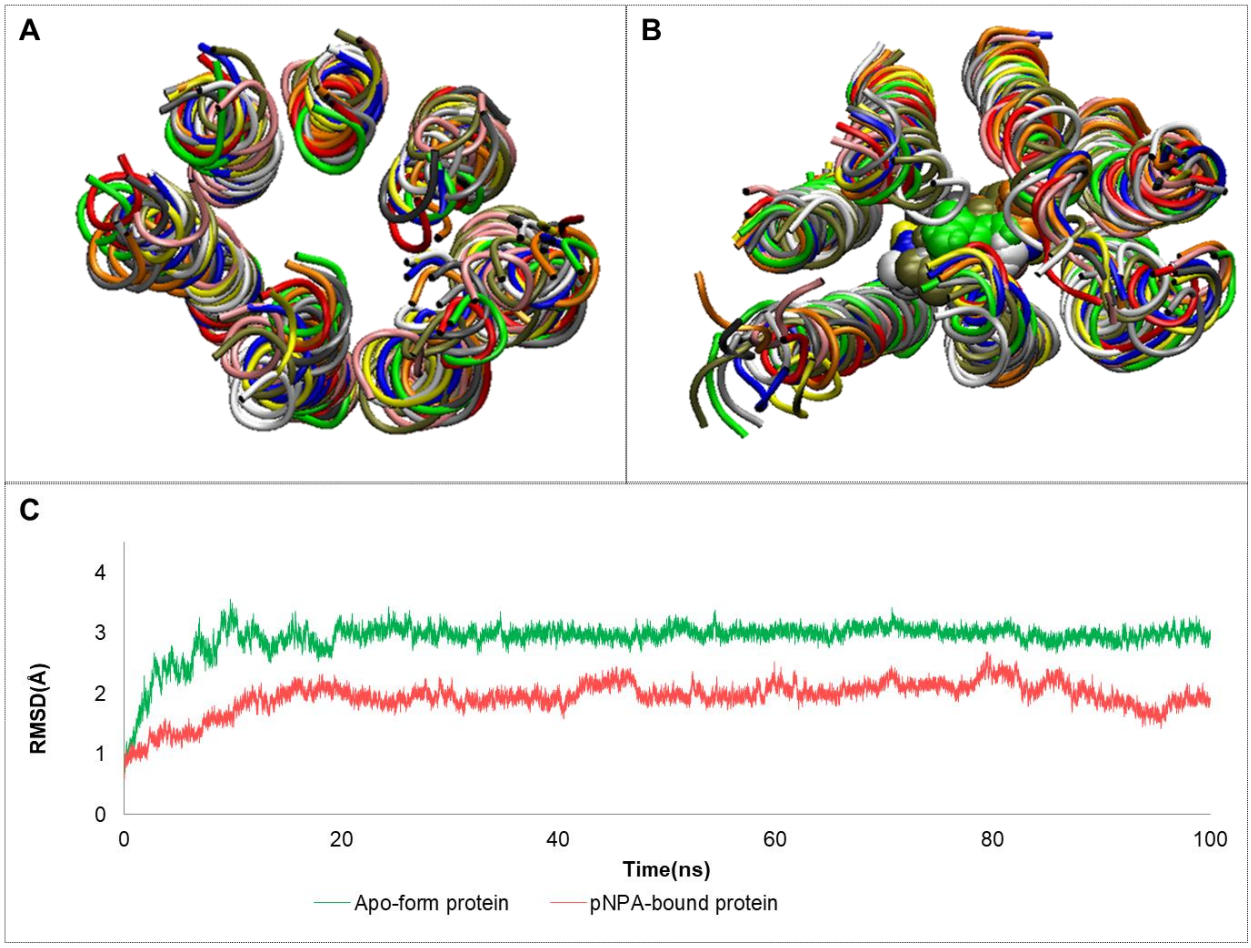


Figure 2. Superposition between 10 representative structures generated by *cpptraj*, after 100 ns of (**A**) the apo-form and (**B**) *p*NPA-bound-protein MDs. Each structure is represented in ribbon form. (**C)** RMSD trends of apo-form (blue line) and *p*NPA-bound protein (red line), calculated on backbone's atoms, expressed in Å.

Specifically, for all the seven α-helices, we observed the high stability except for the α-helical C which reached the average RMSD values >3 Å, after 20 ns of MDs (SI, Figures S1 and S2). MDs also revealed that in the top region, identified by N-terminal protein domain, the side chain and the backbone atoms of Arg30, Leu29, Ala28, Lys27, Glu24 and His22 residues were stabilized by water-bridge interactions, during the entire trajectory. Meanwhile, in the bottom region, turned towards the C-terminal domain, the water-mediated hydrogen bonds were only observed between the side chains of the more hydrophilic residues (Glu2, Lys5 and Arg8).

Afterward, as recently applied, [28] we selected 10 representative structures of 100 ns of unbound protein MDs in order to catch conformational heterogeneity for further evaluation of substrate binding modes into the channel (Figure 2A).

As previously reported[16,43] we used the unnatural substrate *p*NPA for molecular recognition into the channel of the Frankenstein protein, thus generating 10 complexes (Figure S3) for each obtained representative structure. Based on the applied geometrical filters, able to ensure the nucleophilic attack of Cys thiol on the *p*NPA and according to the lower theoretical binding affinity, only one conformation was adopted to predict binding mode. In more detail, the distances of 4.355, 3.795 and 2.815 Å were achieved between S_{Cys} and ester carbonyl moiety of *p*NPA, S_{Cys} and $N_{\varepsilon His}$, $N_{\delta His}$ and O_{Glu} atoms, respectively (Table S2). To confirm that the identified binding pose was stable, other 100 ns of MDs were submitted. Although the average RMSD value of *p*NPA-bound form was 1.939 Å, thus guaranteeing the stability requested by CC-Hept, the RMSD trend for all α-helices revealed that, in the presence of *p*NPA, the protein was associated with the higher plasticity with respect to the unbound protein. Likewise, RMSF analysis confirmed a major flexibility of Frankenstein helices in complex with *p*NPA when compared with those of the apo-form simulation (Figure S4).

The RMSD of *p*NPA was calculated with respect to the protein, after system equilibration, for the evaluation of the substrate's stability. The substrate remained stable for more than half of the MDs, whereupon some sharp peaks revealed the shift of *p*NPA of around 15 Å to the bottom of the channel, and again moved in close proximity to the catalytic triad set for the hydrolytic process (Figure 3).

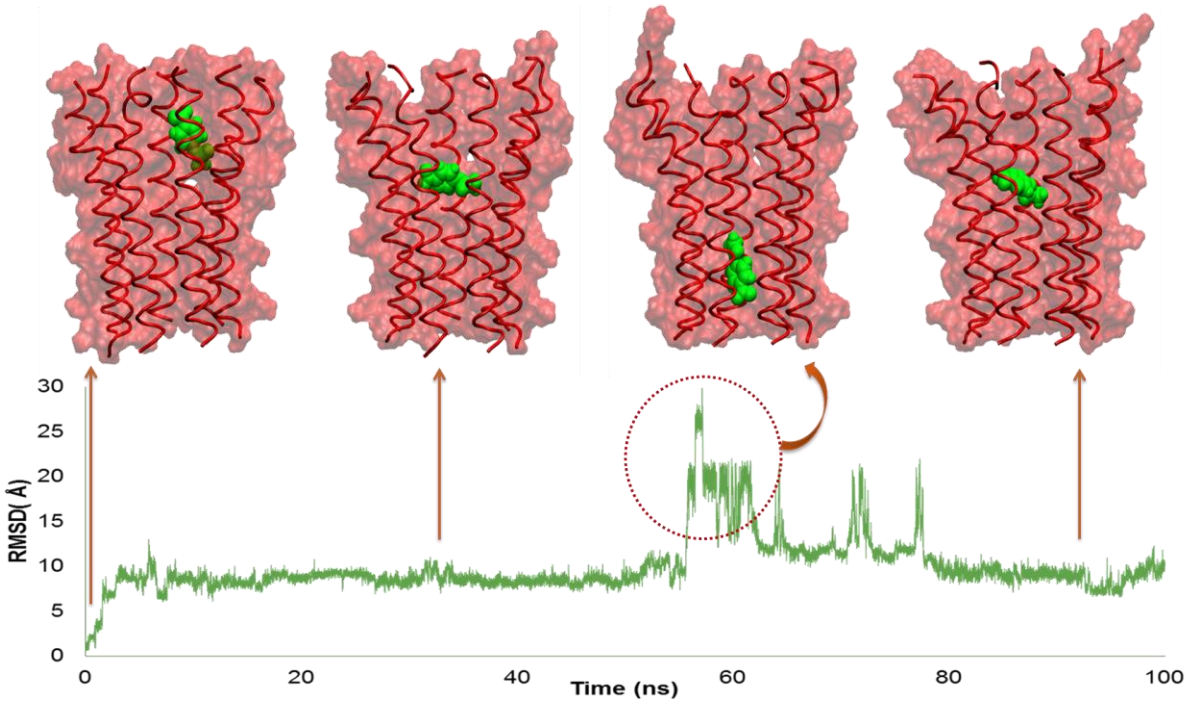


Figure 3. *p*NPA RMSD trend with respect to *de novo* protein during the whole trajectory of MDs. The thick ribbons correspond to pNPA-bound protein every 25 ns of 100 ns of MDs, while pNPA is represented as green CPK.

Analyzing the measure of the channel diameter at the top, middle and bottom sections, we found that the growth to accommodate *p*NPA increased the overall diameter of the α-helical barrel in the top region with respect to that in apo-form protein. Meanwhile, if compared to the initial structure, the average diameter of the region in proximity of the Cys residues, was increased by 0.687 Å and decreased by 1.517 Å for the apo-form and *p*NPA-bound protein respectively, thus suggesting that the required rigidity of the channel was kept in the central region owing to the presence of the substrate (Figure S5).

Exploring the key interactions involved into CC-Hept, we observed that Glu25 engaged hydrogen bond with N_{δ} of His22 with occupancy of 58.34% at an average distance of 2.802 Å. Additionally, an occupancy value of around 80% indicated a water bridge between N_{ε} of His22 and the thiol moiety of Cys_{18}.

After clustering the *p*NPA-bound protein MDs trajectory (Figure 2B) only one structure satisfied the adopted geometrical filters. In this configuration the distances of 4.582, 3.571 and 2.908 Å

were reached for S_{Cys} and ester carbonyl moiety of *p*NPA, S_{Cys} and $N_{\varepsilon His}$, $N_{\delta His}$ and O_{Glu} atoms, respectively (Table S3).

In order to investigate the catalytic mechanism, hybrid QM/MM procedure was applied as previously reported in similar systems.(34) The QM region includes Cys18, His22 and Glu25 residues of the hydrolytic reaction opportunely faced into the hydrophobic channel of CC-Hept and properly aligned to accomplish the catalytic activity *via* three water molecules (w1, w2 and w3) trapped in a hydrogen bonds network both with each other and with the catalytic triad (Figure 1D).

In the proposed mechanism (Figure 4) Cys acts as a nucleophile after its deprotonation by w1 behaving as the proton shuttle to His linked by hydrogen bond to the H^+ final acceptor Glu (activation step described by the TS1). The new-formed active nucleophile Cys-$S^{(-)}$ now can perform the nucleophilic attack on the electrophilic carbon of the reactive ester (*p*NPA) with the synchronous leaving of p-nitrophenolate anion and the formation of a thioester (TS2).

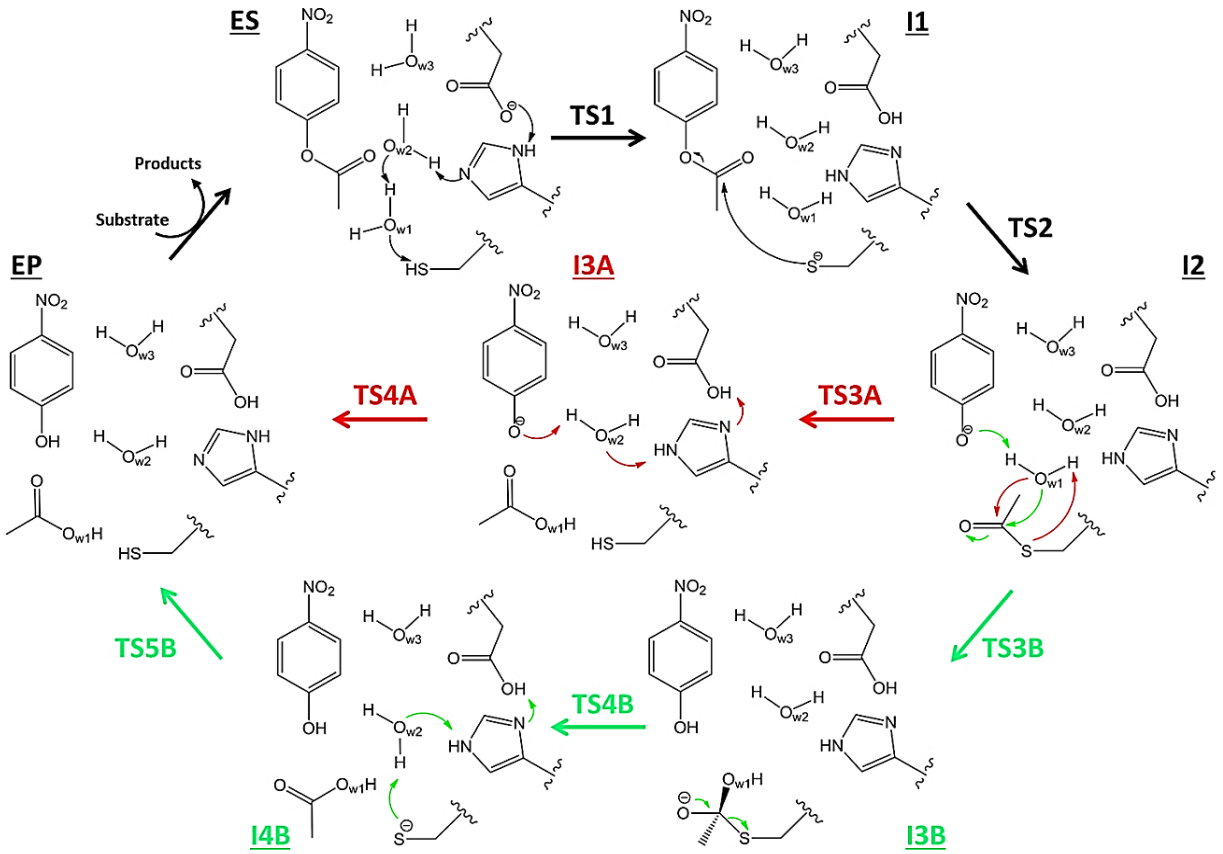


Figure 4. Two proposed reaction mechanisms of *p*NPA hydrolysis catalyzed by *de novo* protein. The achievement of the final EP takes place in two steps for the path A and three steps for the path B, labeled in red and green, respectively.

From the resulting species (I2) the reaction can occur through two different channels (A or B) for affording the final complex (EP). In the channel A, w1 comes into play by nucleophilic attack on the carbon of thioester and simultaneously transfers the proton to sulfur atom generating the S-C bond cleavage (TS3A) with the acetic acid release and –SH moiety restoration on the Cys (I3A). In the last step, His accomplishes the process acting as a general acid/base to deprotonate the Glu residue and to protonate w2 that gives the proton to p-nitrophenolate ion (TS4A). The resulting EP accounts for the regeneration of the catalytic triad.

In the channel B, starting from I2 species three steps are required for obtaining EP. In the first one (occurring *via* TS3B) w1 performs a nucleophilic attack on the electrophile carbon of the

thioester generating a tetrahedral intermediate (I3B). Subsequently, the CH_3COOH formation with consequent C-S bond cleavage (TS4B) occurs giving rise to I4B. Finally, the EP is generated upon restoration of the initial protonation state of the Cys-His-Glu promoted by w2. The computed potential energy surface is reported in Figure 5.

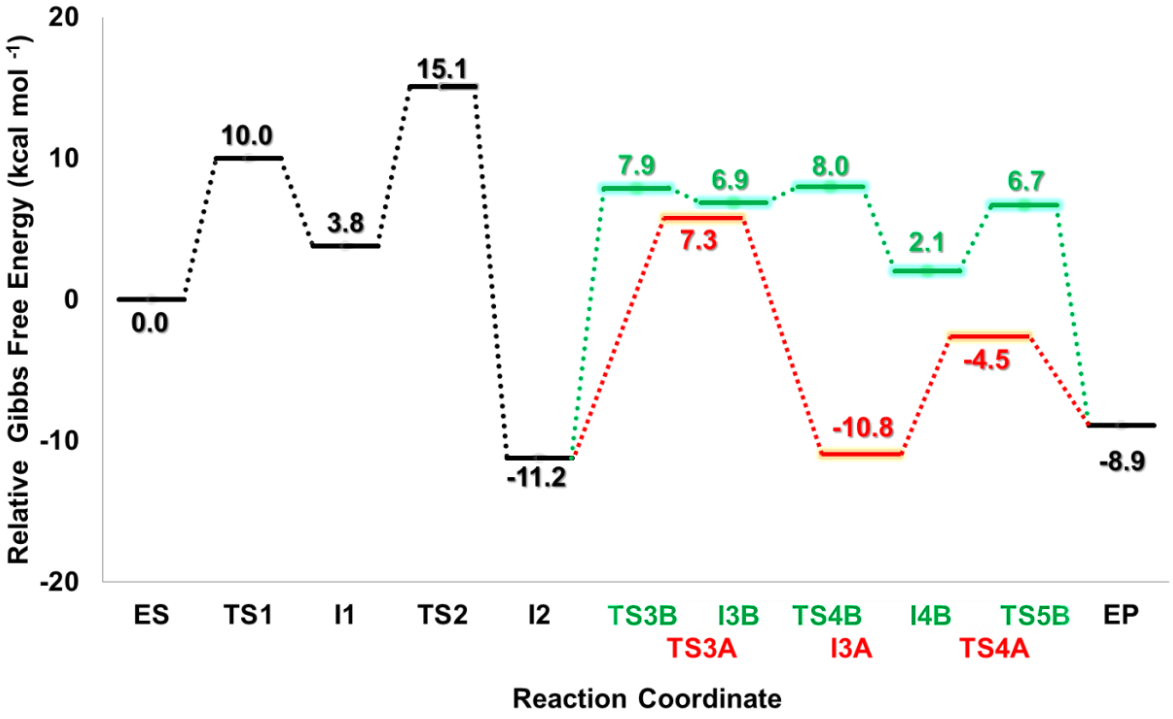


Figure 5. Potential Energy Surfaces for proposed hydrolysis reaction mechanisms catalyzed by *de novo* protein obtained at B3LYP-D3/6-311+(2d,2p):ff99SB//B3LYP/6-31+G(d,p):ff99SB level of theory for A and B channels, in red and green, respectively.

The Cys18-thiol lies at 4.939 Å from the carbonyl of *p*NPA and at 1.959 Å from the w1 that is engaged in hydrogen bond networks with the neighboring w2 and w3 and His22 (Figure 7A). This topology is suitable for the activation process of the thiol group. These three water molecules as hydrogen bond donors/acceptors participate in the chemical steps throughout the entire catalytic cycle (Figure 6).

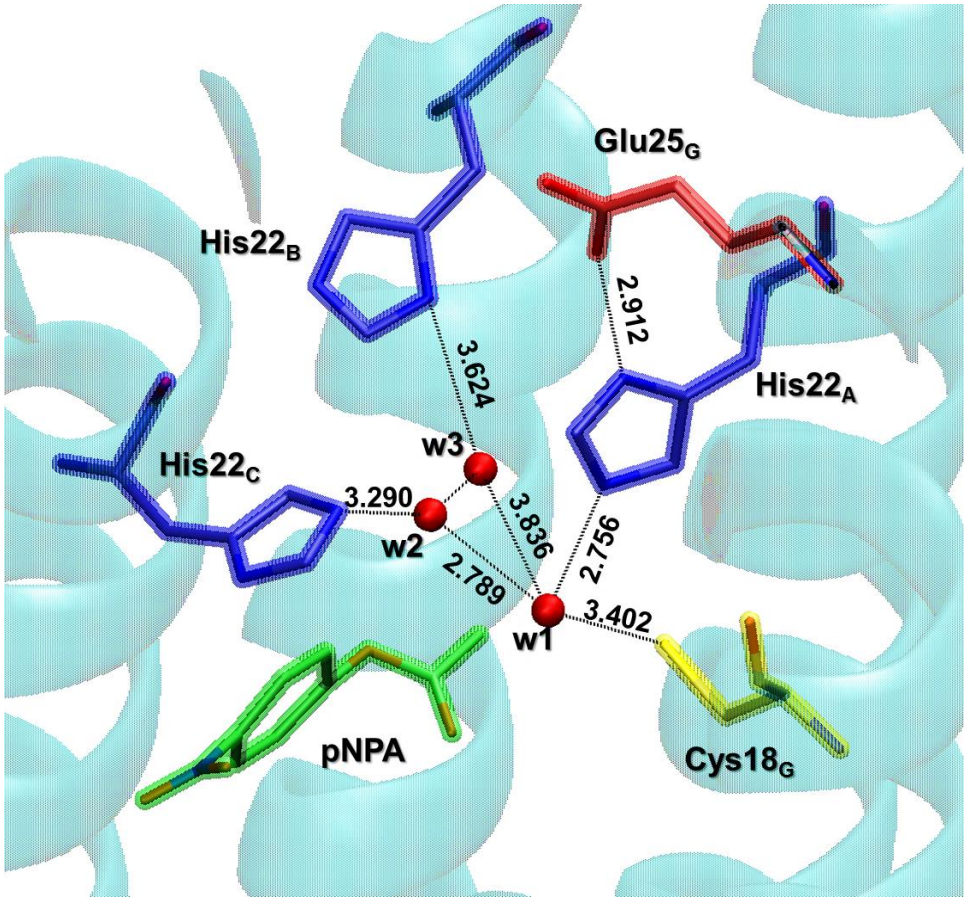


Figure 6. Hydrogen bond network for water molecules explicitly considered in the QMMM model for the *p*NPA-bound protein are represented as black dots. Three water molecules are indicated as red balls. The substrate is shown in green sticks, meanwhile, Cys, His and Glu are indicates as yellow, blue and red sticks, respectively.

As first step, we performed the nucleophilic attack of Cys on *p*NPA with two possible protonation states of the thiol group (Figure 7B). Our calculations showed that, if the thiol group was neutral, the nucleophilic attack did not exhibit transition state and the energy increased throughout the entire linear transit scan (above 40 kcal mol^{-1}), without the formation of any intermediate. On the contrary, the nucleophilic attack with deprotonated thiol moiety leads to a transition state that afforded the thioester species (I2). The deprotonation of the thiol by means of w1 and w2, inserted in the space of 4.715 Å between His and Cys residues, required an activation energy equal to 10.0 kcal mol^{-1} (Figure 5**)**.

Natural charges, arising from the NBO analysis confirm the increased nucleophile character of the soft sulfur atom from the neutral form (-0.138 e) to the deprotonated one (-0.652 e) (Table S4).

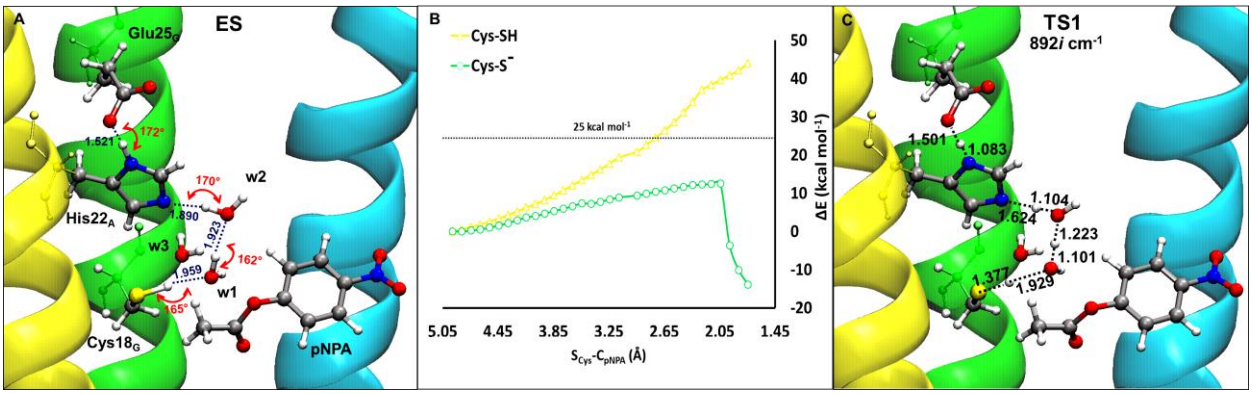


Figure 7. (A) The optimized geometry of ES and respective hydrogen bond network are reported. In details, in blue and red label, the distance and the angle are shown, respectively. (**B**) PES for the nucleophilic attack of the neutral thiol form of Cys-SH and deprotonated one Cys-$S^{(-)}$ to the ester carbonyl. (**C)** TS1 stationary point intercepted along the PES of the hydrolytic process. All the optimizations have been performed by using the B3LYP/6-31+G(d,p):ff99SB level of theory. Main distances (Å) between water molecules and residues are reported in black dots. For clarity, only the amino acid residues taking part in the reaction are shown. Cys18-His22-Glu25 residues and *p*NPA are represented in gray carbon sticks.

The vibrational mode (892*i* cm^{-1}) displayed a linear stretching of the three O_{w1}—H—O_{w2} atoms. In fact, w1 and w2 promote the concerted proton shifts starting from the thiol of Cys until to the oxygen of Glu (Figure 7C). These movements induce a rotation of *p*NPA in the acetyl terminal region that assumes a trans geometry more suitable for the next nucleophile attack. The I1 species, lying at 3.8 kcal mol^{-1}, shows the formed thiolate and the Glu protonation (Figure 8A).

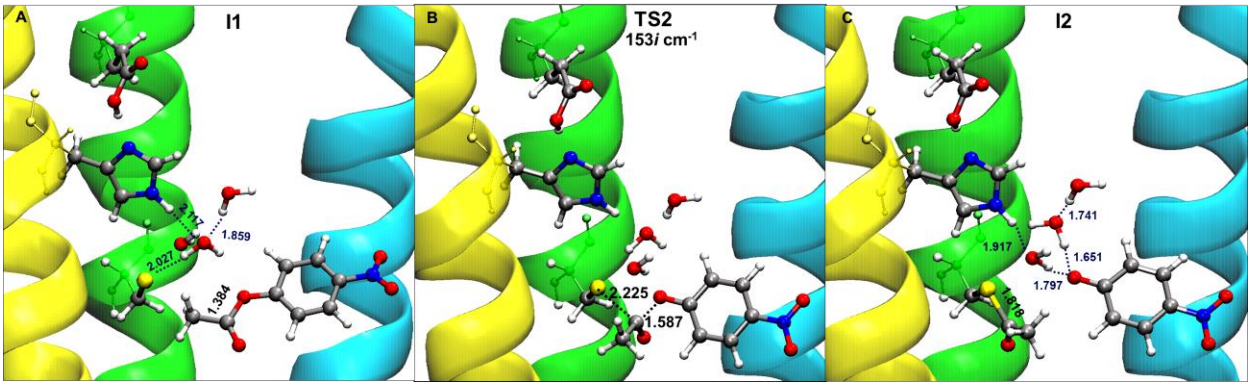


Figure 8. (**A,C**) The optimized geometries of intermediates and (**B**) TS species intercepted along the PES of the hydrolytic process. Main distances (Å) between residues are reported in black dots, meanwhile the water bridges are coloured in blue dots. For clarity, only the amino acid residues taking part in the reaction are shown, in which Cys18-His22-Glu25 residues and *p*NPA are represented in gray carbon sticks.

The TS2 (15.1 kcal mol^{-1} above the ES) clearly depicts the nucleophilic attack by $S^{(-)}$ on the C (2.225 Å) with consequent elongation of the C-O bond (1.529 Å) and the release of the p-nitrophenolate anion (Figure 8B). The related imaginary frequency (153*i* cm^{-1}) well accounts for the S-C and C-O stretching movements. No formation of tetrahedral intermediate takes place. This confirms that the nucleophilic attack occurs simultaneously with the release of the most stable p-nitrophenolate species. The I2 intermediate proposes the acyl intermediate bound to the Cys and is more stable with respect to the starting complex by about 11.2 kcal mol^{-1} (Figure 8C). This strong stabilization explains the experimental isolation and characterization of this intermediate.(15) The next chemical step in the channel A is the hydrolysis of thioester by w1 that donates the proton to the sulfur atom (O_{w1}--H 1.612 Å and H--S 1.780 Å) producing the O_{w1}--C bond formation (1.601 Å) and the S--C bond cleavage (1.977 Å) in concerted way. The corresponding transition state (TS3A) has one imaginary frequency of 752*i* cm^{-1} and a Gibbs activation energy of 18.5 kcal mol^{-1} (Figure 9A).

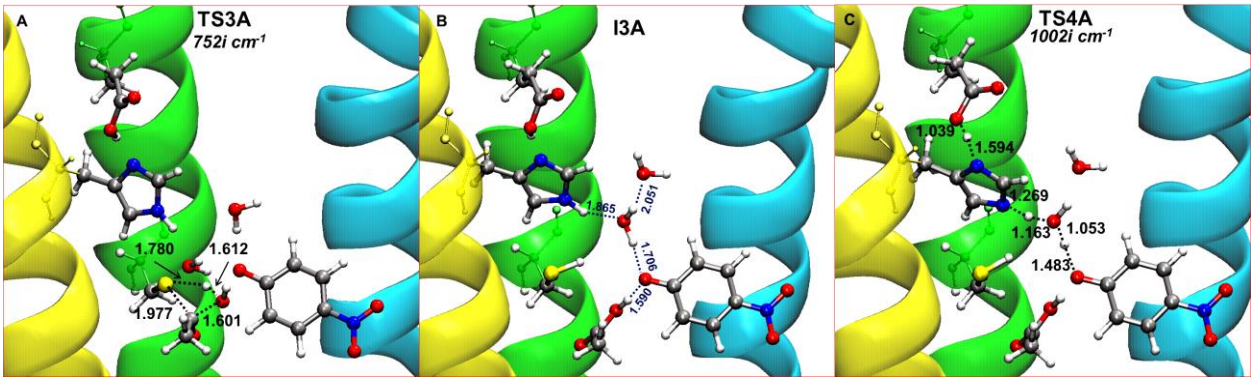


Figure 9. The optimized geometries of (**A**) TS3A, (**B**) I3A and (**C**) TS4A species intercepted along the PES of the hydrolytic process were carried out using B3LYP/6-31+G(d,p):ff99SB level of theory. Main distances (Å) involving water molecules and residues are reported.

The I3A resulting species, almost isoenergetic with I2, must easily restore (6.3 kcal mol^{-1}) the triad along with the protonation of p-nitrophenolate (Figure 9B). This occurs by TS4A where w2 placed between p-nitrophenolate and His simultaneously acts as acid (Ow—H: 1.053 Å and H—O: 1.483 Å) and base (Ow--H: 1.163 Å and H--N: 1.269 Å). The same acid/base role is played by His donating the proton to water and receiving another one from Glu. The imaginary frequency of 1002*i* cm^{-1} confirms the main role played by water in this rearrangement. The reaction Gibbs energy results to be -8.9 kcal mol^{-1} (Figure 9C).

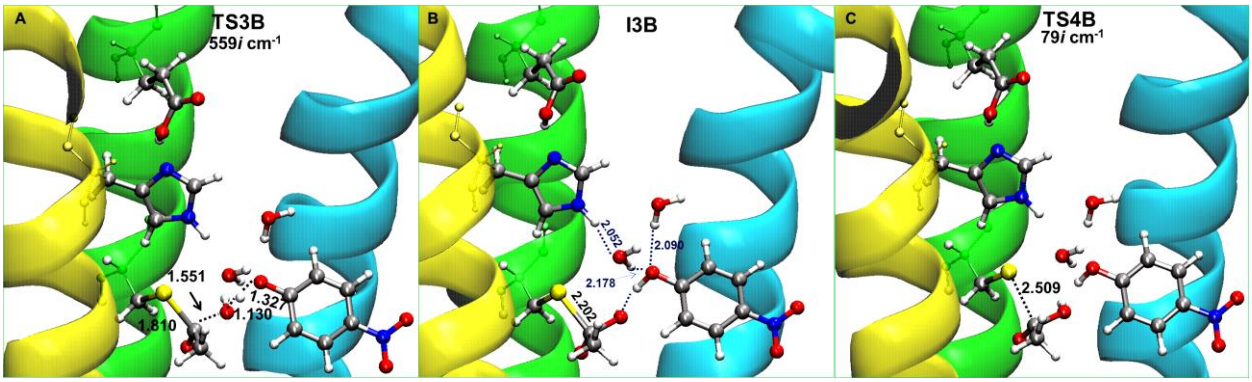


Figure 10. Optimized geometries of (A) TS3B, (B) I3B and (C) TS4B species intercepted along the PES of the hydrolytic process for the channel B. Main distances (Å) involving water molecules and residues are reported in black dots. For clarity, only the amino acid residues

taking part in the reaction are shown, in which Cys_{18}-His_{22}-Glu_{25} residues and *p*NPA are represented in gray carbon sticks.

Starting from I2 (channel B), the reaction proceeds throughout TS3B that describes the protonation of *p*NPA occurring in a concerted way with the formation of tetrahedral intermediate promoted by w1 (O_{w1}—H 1.130 Å and H—O_{pNPA} 1.327 Å, O_{w1}—C 1.551 Å) (Figure 10A). This step requires an amount of 19.1 kcal mol^{-1}. The I3B species reflects the above-explained rearrangement with a more lengthened C—S bond (2.202 Å) (Figure 10B). It evolves most quickly to the end of the reaction by means of TS4B where the C—S bond is practically broken (2.509 Å) (Figure 10C).

In the I4B, lying at 2.1 kcal mol^{-1} above ES, the hydrolysis is accomplished (Figure 11A).

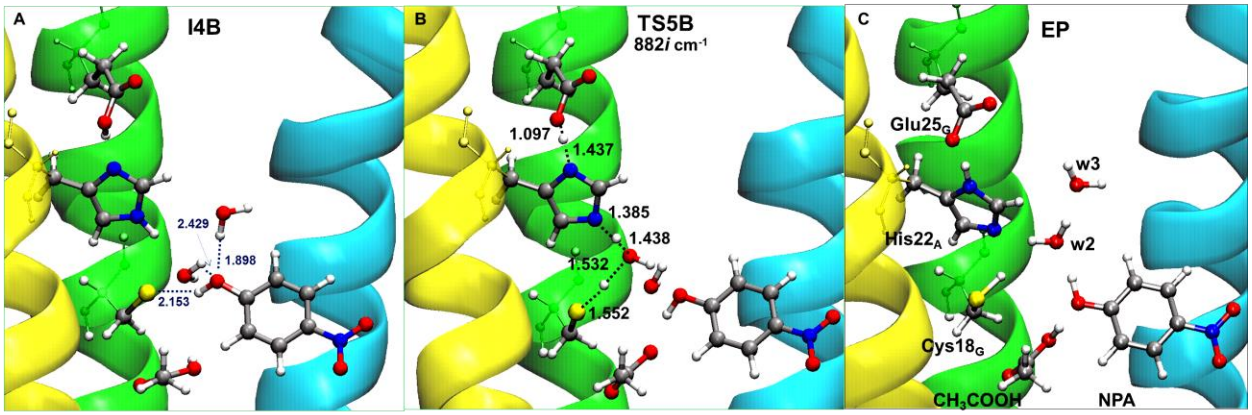


Figure 11. Optimized geometries of (**A**) I4B, (**B**) TS5B and (**C**) EP species intercepted along the PES of the hydrolytic process for the channel B. Main distances (Å) involving water molecules and residues are reported in black dots.

To restore the catalytic cycle multiple proton shifts between S-Cys and imidazole ring of His (O_{w1}—H 1.532 Å and H--S 1.552 Å, O_{w1}—H 1.438 Å and N_{His}—H 1.385 Å, see Figure 11B and 11C) take place with a modest energetic expense (4.6 kcal mol-^{1}).

Discussion

In agreement to the polar/non polar sequence periodicity of the helix, our preliminary structural analysis underlined the stability of CC-Hept-Cys-His-Glu assembly, thus promoting multimeric

association, and installing peculiar active-site residues with hydrolytic activity.[14] MDs revealed that the presence of the *p*NPA led to a top channel growth to accommodate it, if compared to the apo-form protein simulation. Non-wonder, the efficiency of functional Frankenstein lies in this arrangement prone to perform esterase activity into a hydrophobic cavity, ensuring the right balance between plasticity and toughness.[15] The substrate moved in the plane orthogonal to the long axis of the channel and rotated freely.[10] In presence of *p*NPA, the protein showed a slight tightening around catalytic triad, likely due to the increased number of interactions between protein and substrate into the channel promoted also by water molecules present in the catalytic pocket. The presence of water molecules in proximity to the functional catalytic triads during the MDs highlights the effect of the substitution of hydrophobic residues with polar ones into the engineered system. The water molecules (engaged in hydrogen bond networks with each other, with the hydrophilic residues and with the polar moiety of the substrate/product) take part in all the steps of the reaction mechanism.

From the combined QM/MM approach emerges that for both the investigated paths the formation of a tetrahedral intermediate does not occur, since the release of para-nitrophenolate anion takes place simultaneously with the nucleophile attack.[43] Consequently, the first step of the reaction (with the barrier of 10 kcal mol^{-1}) is faster than the resting ones (Figure 5) as in other hydrolytic enzymes.[11]

The thioester species (I2) results to be stabilized enough (11.2 kcal mol^{-1} below the ES) representing the bottleneck of the reaction.

For A and B channels the rate-limiting step (RDS) corresponds to the hydrolysis of the thioester intermediate (18.5 kcal mol^{-1} and 19.1 kcal mol^{-1}, respectively). This result well compares with the barrier obtained by converting the experimental k_{cat} (17-20 kcal mol^{-1}). With respect to the mono-step uncatalyzed reaction (38.6 kcal mol^{-1}) a barrier reduction of about 20 kcal mol^{-1} is observed (Figure S6).

Although the reaction barriers for RDSs are very similar for A and B mechanisms, channel B results more favorable from both thermodynamic and kinetic point of view (Figure 5).

The species intercepted along path B (TS3B, I3B and TS4B) lie in a very flat region of the PES. On the contrary, the corresponding species in the path A (I3A) shows an already restored

nucleophile thiol group with a consequent major stabilization (10.8 kcal mol^{-1} below the ES complex).

Conclusion

In conclusion, our study highlights the origin of the catalytic process of the new engineered hydrolase system. Two reaction mechanism have been analyzed and their feasibility has been established based on calculated energy barriers. The role of the three residues in the active site has been elucidated along the reaction pathways, in which the thiolate moiety of Cys acts as nucleophile. The adopted computational strategy represents an appropriate approach providing the atomistic-level detailed knowledge of the reaction mechanism in the captivating world of *de novo* protein design. Furthermore, it contributes to corroborate and validate experimental findings on the recently catalytically active Cys-His-Glu triad installed into engineered protein frameworks.

Acknowledgements

Financial support from the Università degli Studi della Calabria-Dipartimento di Chimica e Tecnologie Chimiche (CTC) is acknowledged.

References

1. Galante, Y. M.; Formantici, C. *Current organic chemistry* **2003,** *7* (13), 1399.
2. Walker, J. M., The bicinchoninic acid (BCA) assay for protein quantitation. In *The Protein Protocols Handbook*, Springer: 2002; pp 11-14.
3. Woolfson, D. N.; Bartlett, G. J.; Burton, A. J.; Heal, J. W.; Niitsu, A.; Thomson, A. R.; Wood, C. W. *Curr. Op.Struct. Bio.* **2015,** *33*, 16.
4. Schmid, A.; Dordick, J.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001,** *409* (6817), 258.
5. Bozhüyük, K. A.; Fleischhacker, F.; Linck, A.; Wesche, F.; Tietze, A.; Niesert, C.-P.; Bode, H. B. *Nat. Chem.* **2018,** *10* (3), 275.

6. Reig, A. J.; Pires, M. M.; Snyder, R. A.; Wu, Y.; Jo, H.; Kulp, D. W.; Butch, S. E.; Calhoun, J. R.; Szyperski, T.; Solomon, E. I. *Nat. Chem.* **2012,** *4* (11), 900.
7. Polizzi, N. F.; Wu, Y.; Lemmin, T.; Maxwell, A. M.; Zhang, S.-Q.; Rawson, J.; Beratan, D. N.; Therien, M. J.; DeGrado, W. F. *Nat. Chem.* **2017,** *9* (12), 1157.
8. Huang, P.-S.; Oberdorfer, G.; Xu, C.; Pei, X. Y.; Nannenga, B. L.; Rogers, J. M.; DiMaio, F.; Gonen, T.; Luisi, B.; Baker, D. *Science* **2014,** *346* (6208), 481.
9. Woolfson, D. N., The design of coiled-coil structures and assemblies. In *Advances in protein chemistry*, Elsevier: 2005; Vol. 70, pp 79-112.
10. Thomas, F.; Dawson, W. M.; Lang, E. J.; Burton, A. J.; Bartlett, G. J.; Rhys, G. G.; Mulholland, A. J.; Woolfson, D. N. *ACS Synth. Biol.* **2018,** *7* (7), 1808.
11. Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A. *J. Am. Chem. Soc.* **2012,** *134* (39), 16197.
12. Koshland Jr, D. E. *Angew. Chem. Int. Ed. in English* **1995,** *33* (23-24), 2375.
13. Radkiewicz, J. L.; Brooks, C. L. *J. Am. Chem. Soc.* **2000,** *122* (2), 225.
14. Smith, A. J.; Müller, R.; Toscano, M. D.; Kast, P.; Hellinga, H. W.; Hilvert, D.; Houk, K. *J. Am. Chem. Soc.* **2008,** *130* (46), 15361.
15. Burton, A. J.; Thomson, A. R.; Dawson, W. M.; Brady, R. L.; Woolfson, D. N. *Nat. Chem.* **2016,** *8* (9), 837.
16. Makhlynets, O. V.; Korendovych, I. V. *Nat. Chem.* **2016,** *8* (9), 823.
17. Der, B. S.; Edwards, D. R.; Kuhlman, B. *Biochemistry* **2012,** *51* (18), 3933.
18. Ma, S.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2007,** *129* (44), 13633.
19. Case, D. A.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.;Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.;LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. (2017), AMBER 2017, University of California, San Francisco.
20. Schafmeister, C.; Ross, W.; Romanovski, V., LEaP. *University of California, San Francisco* **1995**.

21. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comp.* **2015,** *11* (8), 3696.
22. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. *Nucleic acids research* **2012,** *40* (W1), W537.
23. Laurie, A. T.; Jackson, R. M *Bioinformatics* **2005,** *21* (9), 1908.
24. Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. *J.Chem. Phys.* **1984,** *81* (8), 3684.
25. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys* **1993,** *98* (12), 10089.
26. Case, D.; Darden, T.; Cheatham III, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K., AmberTools 16, University of California, San Francisco, 2016.
27. Amaro, R. E.; Baron, R.; McCammon, J. A *J. Computer-Aided Mol. Design* **2008,** *22* (9), 693.
28. Romeo, I.; Marascio, N.; Pavia, G.; Talarico, C.; Costa, G.; Alcaro, S.; Artese, A.; Torti, C.; Liberto, M. C.; Focà, A. *Chem. Sel.* **2018,** *3* (21), 6009.
29. Baron, R.; McCammon, J. A *Biochemistry* **2007,** *46* (37), 10629.
30. Trott, O.; Olson, A. J. *J. Comp. Chem.* **2010,** *31* (2), 455.
31. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A *J. Comp.Chem.* **2004,** *25* (9), 1157.
32. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993,** *97* (40), 10269.
33. Svensson, M.; Humbel, S.; Froese, R. D.; Matsubara, T.; Sieber, S.; Morokuma, K. *The J. Phys. Chem.* **1996,** *100* (50), 19357.
34. a) Prejanò, M.; Marino, T.; Russo, N. *Frontiers in chemistry* **2018,** *6*, 249. b) Piazzetta, P.; Marino, T.; Russo, N. *Inor. Chem.* **2014,** *53* (7), 3488. c) Sousa, S. F.; Ribeiro, A. J.; Neves, R. P.; Brás, N. F.; Cerqueira, N. M.; Fernandes, P. A.; Ramos, M. J. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1281.
35. Piazzetta, P.; Marino, T.; Russo, N. *Phys. Chem. Chem. Phys.* **2015,** *17* (22), 14843.
36. Frisch, M. J.; G. W., Trucks, Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato,

M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2013

37. Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery Jr, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006,** *2* (3), 815.
38. Becke, A. D. *J. Chem. Phys.* **1993,** *98* (7), 5648.
39. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37* (2), 785.
40. Blomberg, M. R.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. *Chem. Rev.* **2014,** *114* (7), 3601.
41. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comp. Chem.* **2011,** *32* (7), 1456.
42. E. D. Glendening, A. E. Reed, J. E. Carpenter,F.Weinhold, NBO, version 3.1.
43. Menger, F.; Ladika, M. *J. Am. Chem. Soc.* **1987,** *109* (10), 3145.
44. Fuhrmann, C. N.; Daugherty, M. D.; Agard, D. A. *J. Am. Chem. Soc.* **2006,** *128* (28), 9086.

Mechanistic insights of hydrolytic activity into a *de novo* Functional protein framework.

Mario Prejano', Isabella Romeo, Tiziana Marino*, Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, 87036 Rende, CS, Italy.

correspondence should be addressed to: tiziana.marino65@unical.it

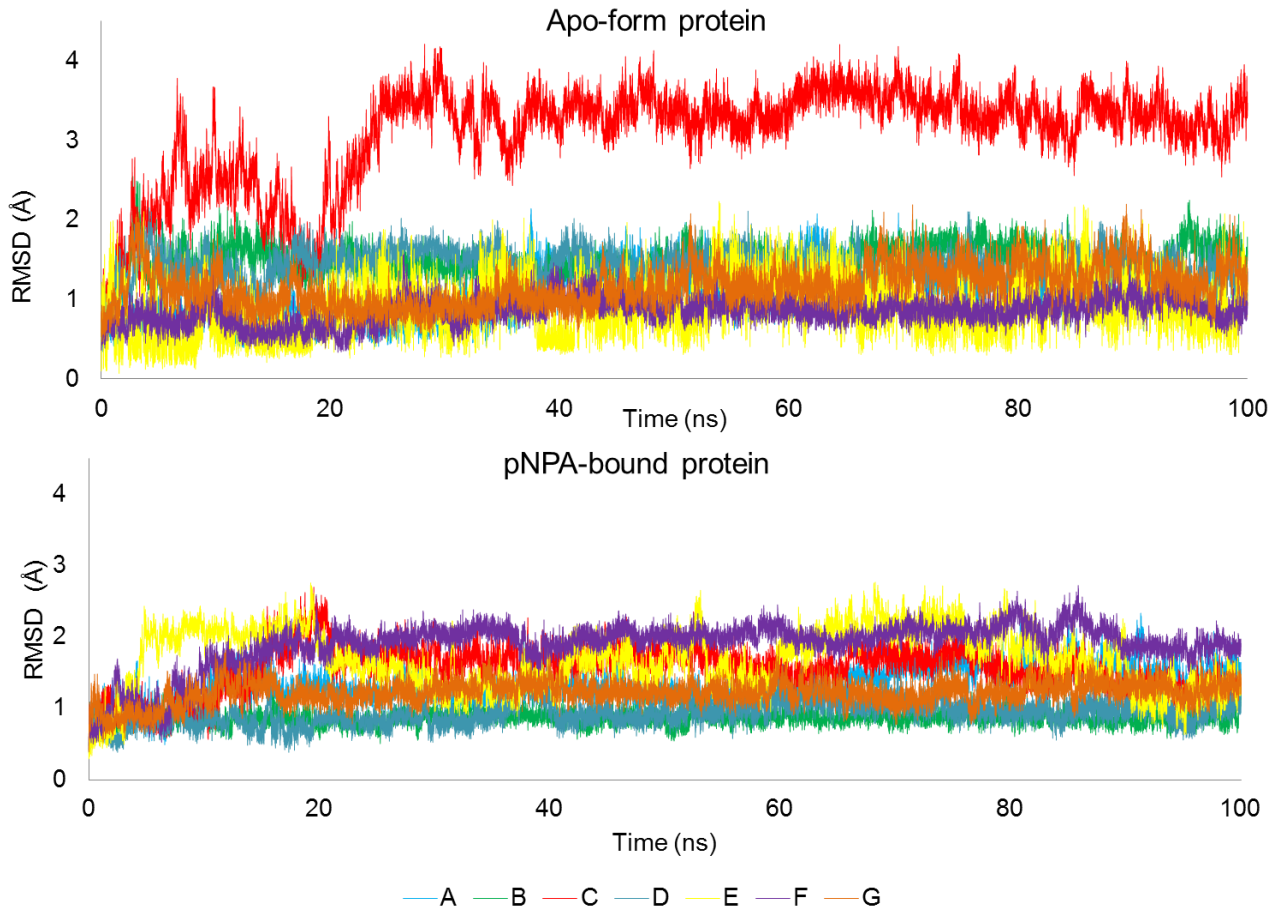


Figure S1. Plot of the RMSD trend, calculated on the seven α-helices backbone atoms with respect to the initial structure, of the apo-form (top) and *p*NPA-bound protein (bottom), during 100 ns of MDs. All seven chains, containing each one 1-30 residues, are indicated by A-G capital letters.

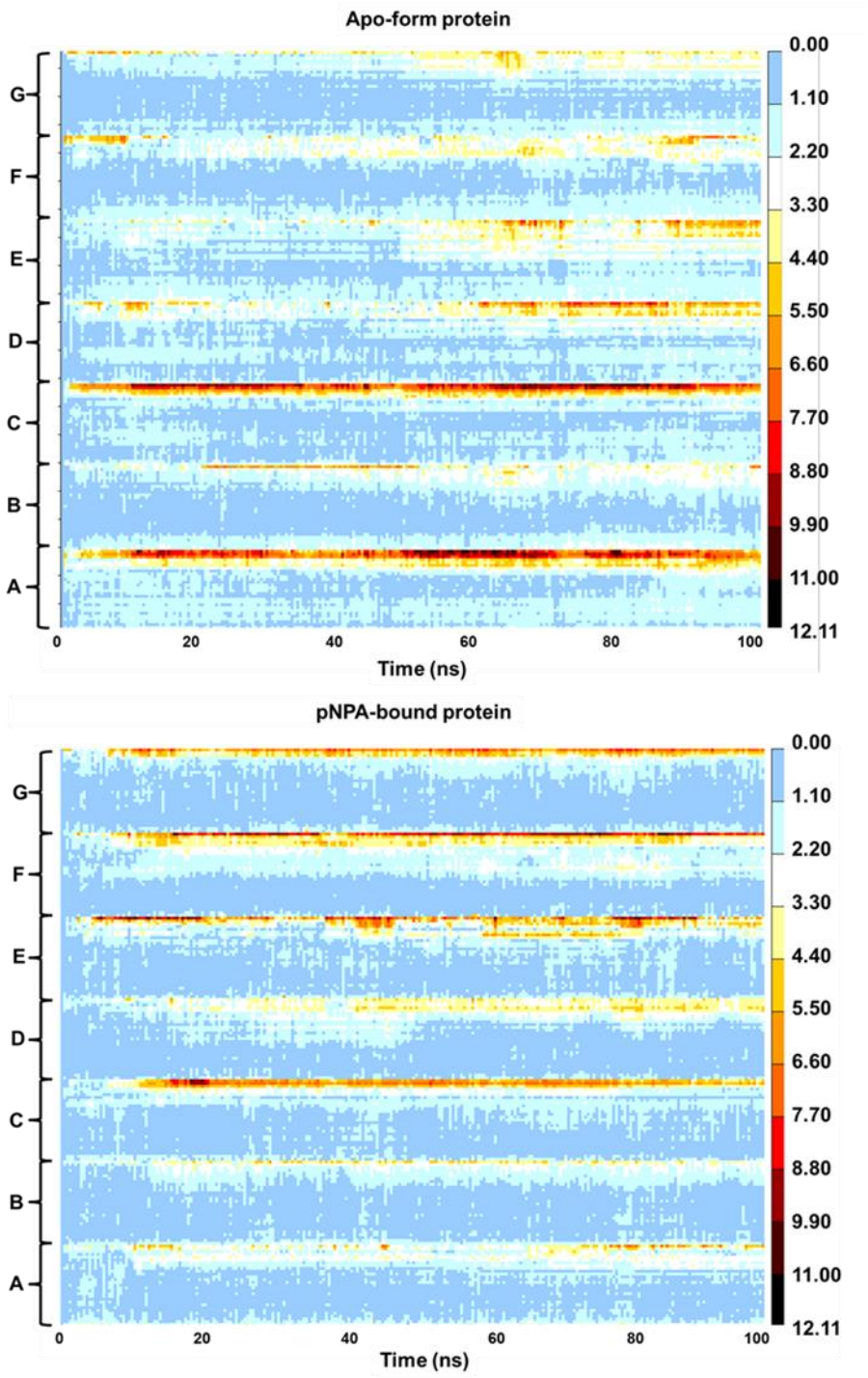


Figure S2. 2D-RMSD of all residues of *de novo* protein as a function of time during 100 ns of MDs in apo-form (top) and *p*NPA-bound protein (bottom). All seven chains, containing each one 1-30 residues, are indicated by A-G capital letters.

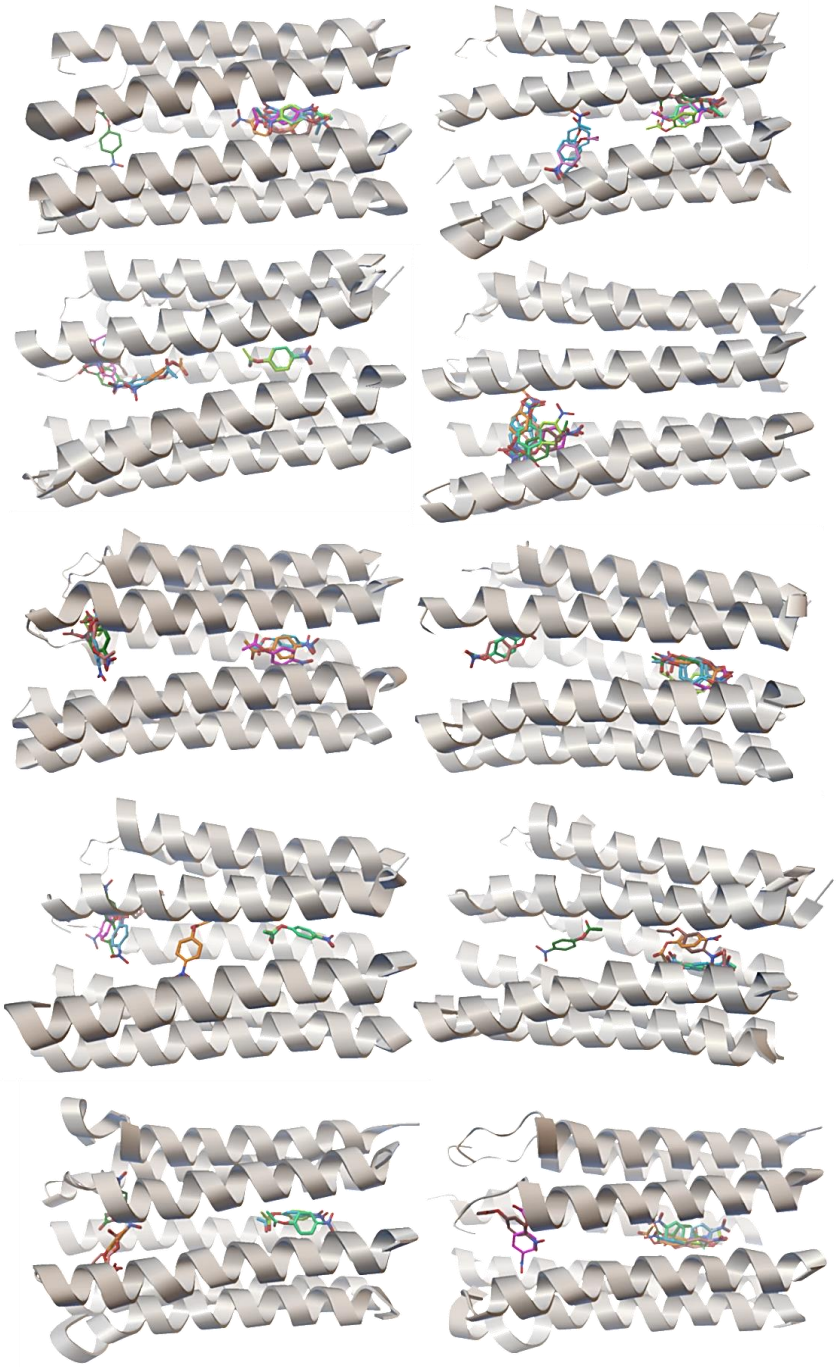


Figure S3. Representation of the unnatural substrate *p*NPA binding mode into the channel of the Frankenstein protein for each obtained representative structure. 10 poses are generated by using AutoDock Vina. The protein and the relative *p*NPA binding poses are shown in light grey cartoon and coloured carbon sticks, respectively.

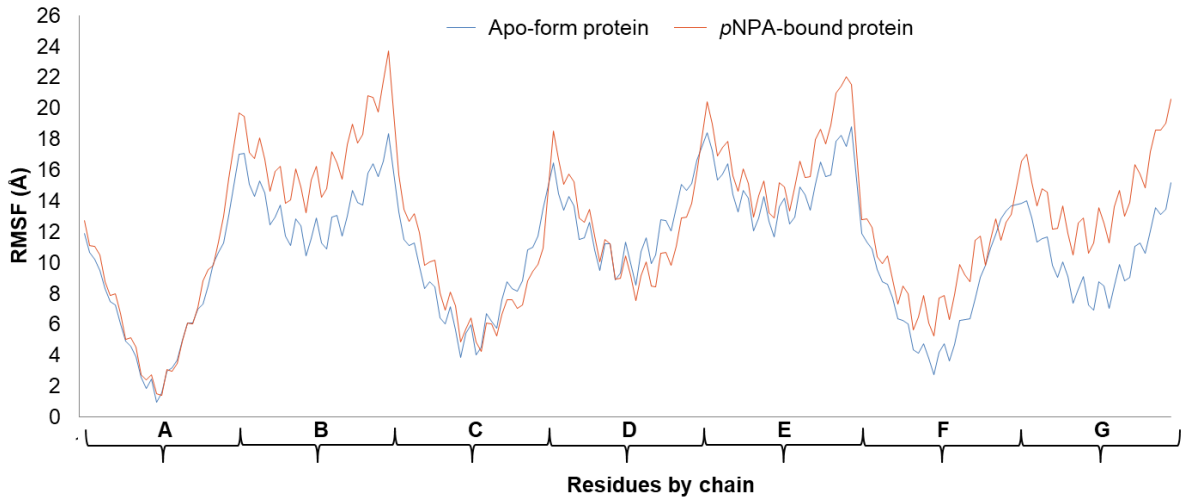


Figure S4. RMSF trends, calculated on all atoms, of apo-form and *p*NPA-bound protein, in blue and red lines, respectively. All seven chains, containing each one 1-30 residues, are indicated by A-G capital letters.

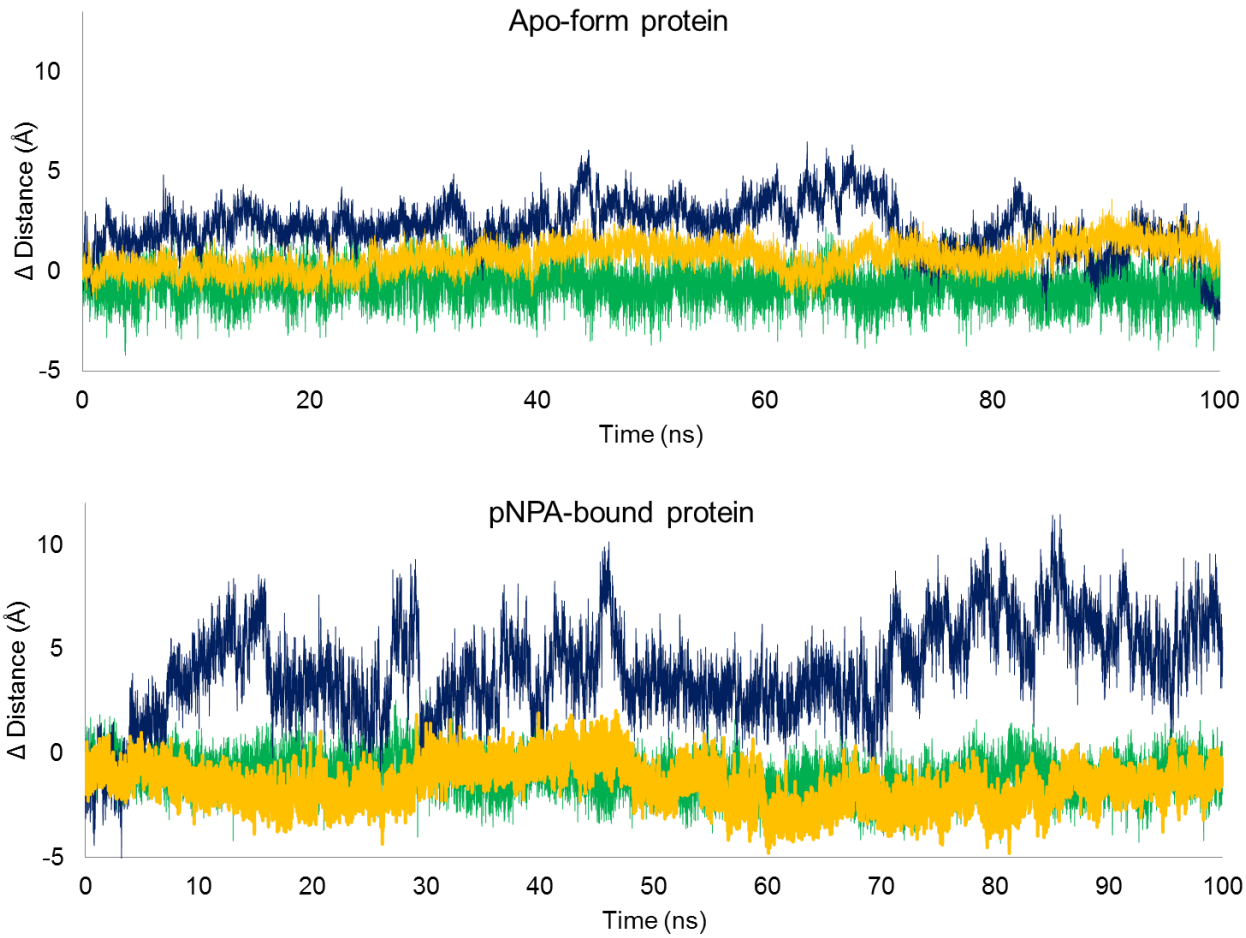


Figure S5. Plot of the distance calculated between the centroids of all the seven Gly_1, Cys_{18} and Arg_{30} residues of the protein and the corresponding Cα of Gly, Cys and Arg of each chain as a measure of the diameter of the channel at bottom (green line), middle and top (yellow and blue lines), respectively, relative to the initial structure. The time of simulations (ns) and the values of the Cα distance (Å) are reported, respectively, on the abscissa and on the ordinate axes for apo-form and *p*NPA-bound protein.

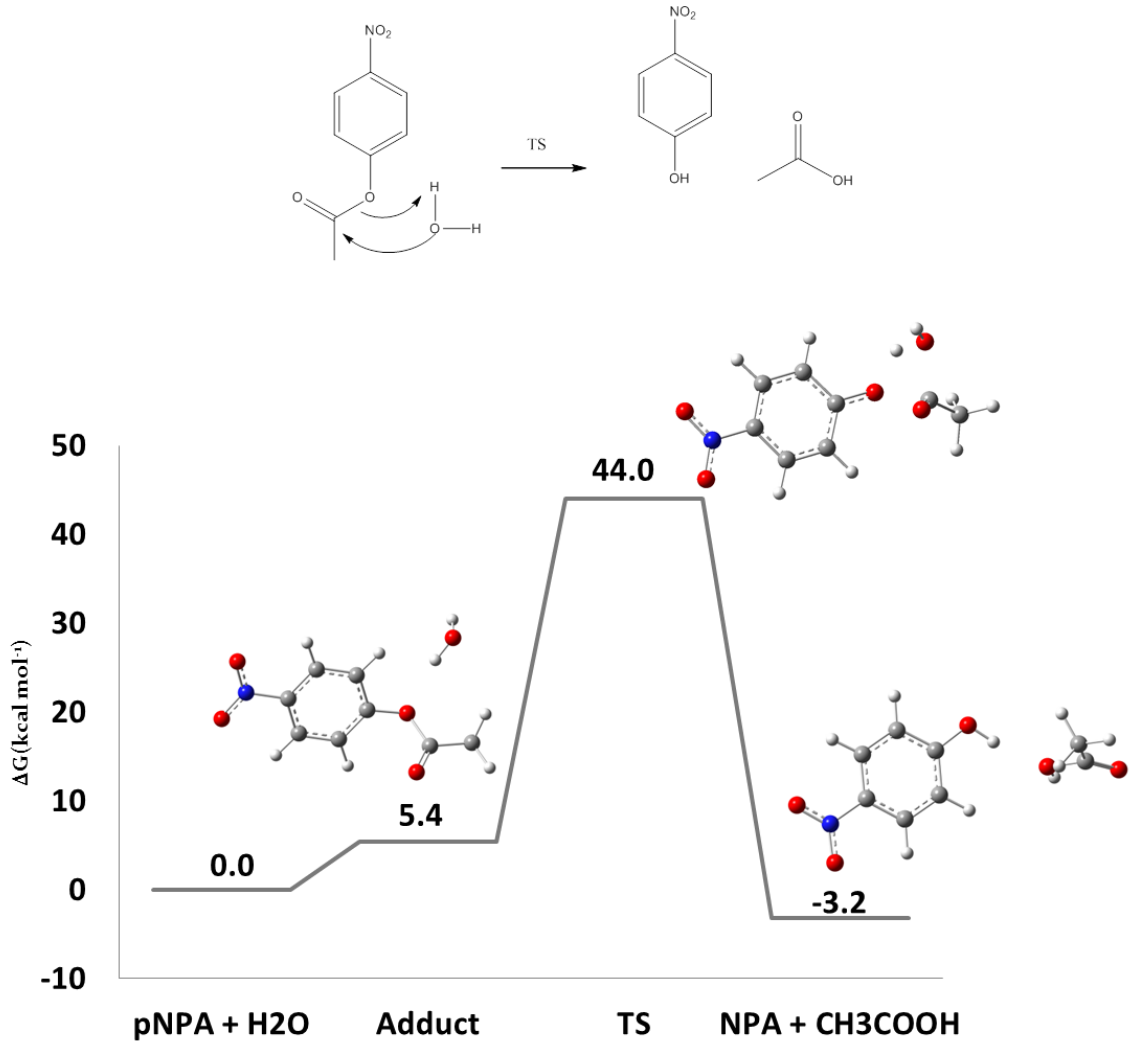
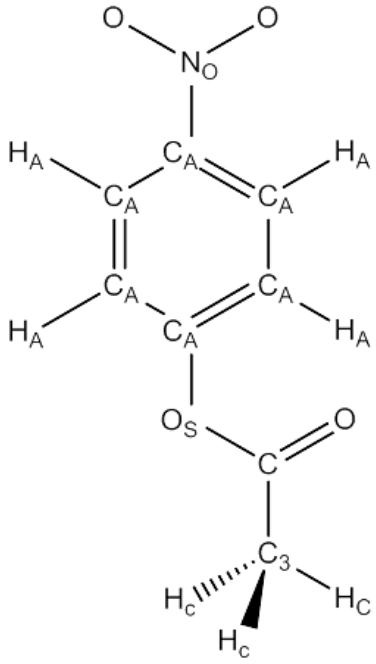


Figure S6. One step uncatalyzed reaction involving *p*NPA and one water molecule, calculated at B3LYP-D3/6-31+G(d,p) level of theory.

Table S1. Calculated parameters for *p*NPA.



Atomtype	Charge
Ca	-0.010
Ha	0.193
No	0.813
O	-0.473
Os	-0.524
C	0.938
c3	-0.558
Hc	0.162

Bond	K_i / kcal mol^{-1} Å^{-2}	l_0/Å
ca-ca	461.10	1.398
ca-no	321.70	1.469
ca-ha	345.80	1.086
no-o	741.80	1.226
ca-os	376.60	1.370
os-c	390.80	1.358
c -c3	313.00	1.524
c –o	637.70	1.218
c3-hc	330.60	1.097

Angle	K_{an}/ Kcal mol^{-1} rad^{-2}	θ_0/deg
ca-ca-ca	66.620	120.020

ca-ca-ha	48.180	119.880
ca-no-o	68.700	117.760
ca-ca-no	66.770	119.010
ca-ca-os	69.580	119.200
o -no-o	76.730	125.080
ca-os-c	63.410	121.150
os-c -c3	68.890	110.720
os-c -o	75.320	123.250
c -c3-hc	46.930	108.770
c3-c -o	67.400	123.200
hc-c3-hc	39.400	107.580

Dihedral	d	$V_n/2$ Kcal mol^{-1} rad^{-2}	γ/deg	n
ca-ca-ca-ha	1	3.625	180.000	2.000
ca-ca-ca-ca	1	3.625	180.000	2.000
ca-ca-no-o	1	0.600	180.000	2.000
ca-ca-ca-os	1	3.625	180.000	2.000
ca-ca-ca-no	1	3.625	180.000	2.000
ca-ca-os-c	1	0.900	180.000	2.000
ha-ca-ca-ha	1	3.625	180.000	2.000
ha-ca-ca-os	1	3.625	180.000	2.000
ha-ca-ca-no	1	3.625	180.000	2.000
ca-os-c -c3	1	2.700	180.000	2.000
ca-os-c -o	1	2.700	180.000	2.000
os-c -c3-hc	1	0.000	180.000	2.000

hc-c3-c -o	1	0.800	0.000	-1.000
hc-c3-c -o	1	0.000	0.000	-2.000
hc-c3-c -o	1	0.080	180.000	3.000

Improper	V_n/ Kcal rad^{-2}	mol^{-1}	γ/deg	n
ca-ca-ca-no	1.1		180.0	2.0
ca-ca-ca-ha	1.1		180.0	2.0
ca-o -no-o	1.1		180.0	2.0
ca-ca-ca-os	1.1		180.0	2.0
c3-o -c -os	10.5		180.0	2.0

VdW	r_0 /Å	ε/ kcal mol^{-1}
ca	1.908	0.0860
ha	1.459	0.0150
no	1.824	0.1700
o	1.6612	0.2100
os	1.6837	0.1700
c	1.908	0.0860
c3	1.908	0.1094
hc	1.487	0.0157

Table S2. Best *p*NPA docked pose for each representative structure and adopted geometrical parameters in order to predict binding mode into *de novo* protein.

	Vina binding scores (kcal/mol)	S_{Cys} - C_{pNPA} atomic distance (Å)	S_{Cys} – $N_{\varepsilon His}$ atomic distance (Å)	$N_{\delta His}$ – O_{Glu} Atomic distance (Å)
Cluster 1	-5.4	13.697	3.472	8.584
Cluster 2	-5.5	4.355	3.795	2.815
Cluster 3	-5.7	13.487	5.732	12.688
Cluster 4	-6.3	5.796	6.306	2.855
Cluster 5	-5.8	11.351	3.357	9.517
Cluster 6	-5.0	11.200	6.098	6.066
Cluster 7	-5.3	13.027	5.273	11.417
Cluster 8	-5.3	4.566	5.356	9.883
Cluster 9	-5.2	11.066	3.553	8.030
Cluster 10	-4.9	7.732	7.963	5.774

Table S3. Adopted geometrical filters on each representative structure of *p*NPA-bound protein in order to select the starting point for the QMMM investigation (Cluster 5).

	S_{Cys} - C_{pNPA} atomic distance (Å)	S_{Cys} – $N_{\varepsilon His}$ atomic distance (Å)	$N_{\delta His}$ – O_{Glu} atomic distance (Å)
Cluster 1	5.011	7.521	9.165
Cluster 2	6.445	6.184	6.294
Cluster 3	4.831	7.322	2.902
Cluster 4	4.981	4.000	5.920
Cluster 5	4.582	3.571	2.908
Cluster 6	3.942	3.444	9.470
Cluster 7	4.683	6.093	2.782
Cluster 8	6.581	4.771	2.951
Cluster 9	3.941	7.512	4.285
Cluster 10	4.151	6.385	6.252

Table S4. NBO charges (|e|) of selected atoms in all the species found on the PES.

	ES	TS1	I1	TS2	I2
S_{Cys}	-0,138	-0,166	-0,652	-0,239	0,174
C_{NPA}	0,837	0,841	0,820	0,616	0,401
O_{w1}	-1,013	-1,101	-1,090	-1,066	-1,041
$O1_{NPA}$	-0,593	-0,599	-0,603	-0,667	-0,577
$O2_{NPA}$	-0,559	-0,563	-0,535	-0,604	-0,809
	TS3A	I3A	TS4A	EP	
S_{Cys}	-0,138	-0,166	-0,652	-0,239	
C_{NPA}	0,837	0,841	0,820	0,616	
O_{w1}	-1,013	-1,101	-1,090	-1,066	
$O1_{NPA}$	-0,593	-0,599	-0,603	-0,667	
$O2_{NPA}$	-0,559	-0,563	-0,535	-0,604	
	TS3B	I3B	TS4B	I4B	TS5B
S_{Cys}	0,080	-0,261	-0,501	-0,715	-0,273
C_{NPA}	0,402	0,584	0,699	0,819	0,820
O_{w1}	-0,838	-0,850	-0,816	-0,747	-0,769
$O1_{NPA}$	-0,732	-0,749	-0,681	-0,615	-0,693
$O2_{NPA}$	-0,842	-0,788	-0,755	-0,789	-0,748

Table S5. Energy contributions, calculated at ONIOM[B3LYP-D3/6-311+G(2d,2p):ff99SB] level of theory, extracted to obtain Potential Energy Surfaces, for each stationary point.

Species	E_{B3LYP} (a.u.)	E_{ZPE} (a.u.)	E_{D3} (a.u.)	E_{total} (a.u)
ES	-6065.276575	29.114817	-0.177247	-6036.339006
TS1	-6065.250377	29.104580	-0.177202	-6036.322998
I1	-6065.266732	29.112009	-0.178182	-6036.332906
TS2	-6065.256521	29.115570	-0.174065	-6036.315019
I2	-6065.304601	29.119100	-0.171312	-6036.356814
TS3A	-6065.270381	29.117517	-0.174539	-6036.327403
TS3B	-6065.266646	29.112971	-0.172782	-6036.326458
I3A	-6065.300779	29.115517	-0.170976	-6036.356239
I3B	-6065.273341	29.116018	-0.170753	-6036.328076
TS4A	-6065.284532	29.107267	-0.168867	-6036.346132
TS4B	-6065.270212	29.115973	-0.172060	-6036.326298
I4B	-6065.279843	29.116563	-0.172460	-6036.335741
TS5B	-6065.271466	29.106182	-0.163046	-6036.328330
EP	-6065.298616	29.114800	-0.169381	-6036.353198

GIUDIZIO FINALE

Giudizio del Supervisore sull'attività di ricerca del dottorando Mario Prejanò - Dottorato di Ricerca in Medicina Traslazionale (XXXI Ciclo)

Il Dott. Mario Prejanò ha iniziato il giorno 01/11/2015 il corso di dottorato di ricerca in Medicina Traslazionale (Indirizzo: Progettazione Molecolare) con un progetto di ricerca avente come oggetto l'investigazione teorica di importanti reazioni catalizzate da enzimi afferenti a diverse classi enzimatiche e di cui alcuni dotati di un potere catalitico poco usuale. Durante i tre anni di attività, in particolare, il dottorando si è dedicato allo studio di aspetti principali della catalisi promossa da enzimi contenenti cofattori inorganici e/o organici coinvolti in importanti processi di interesse biologico. Gli studi sono stati condotti mediante l'utilizzo di avanzate metodiche quantomeccaniche basate sulla teoria del funzionale della densità. Il Dottorando ha avuto modo di apprendere i principali criteri per la scelta oculata del protocollo computazionale, dell'approccio per la simulazione dell'effetto del solvente nei vari casi esaminati.

In particolare, sono stati studiati :

a. Il ruolo di ioni metallici diversi nella stessa tasca catalitica di enzimi (*metanolo deidrogenasi e nitrile idratasi*) è stato esplorato applicando una tecnica di indagine teorica interamente condotta a livello quanto-meccanico "QM cluster approach". Tale studio è stato indispensabile per contribuire a far luce sul ruolo degli ioni metallici (Ca/Ce per la metanolo deidrogenasi e Fe/Co per la nitrile idratasi) analizzando i passaggi elementari del meccanismo catalitico. Un altro importante aspetto della catalisi esaminato, utilizzando lo stesso metodo "QM cluster approach", è quello inerente all'attività catalitica promossa da zinco enzimi contenenti residui amminoacidici modificati nel sito catalitico (*peptidoglicano N-deacetilasi*). Il meccanismo catalitico dell'ammido idrolasi, *LigW decarbossilasi*, contenente manganese nel sito attivo, è stato ampiamente esaminato applicando diverse tecniche di indagine teorica (QM, QM/MM e MD).
b. L'origine del potere catalitico degli enzimi legato non solo alla stabilizzazione degli stati di transizione che determinano la velocità della reazione, ma anche alla "destabilizzazione" di intermedi per mantenere le barriere dei passaggi successivi

entro un'altezza accettabile è stata investigata nell'enzima transchetolasi, *hTK*, contenente un cofattore di natura organica. A tale scopo l'indagine è stata eseguita attraverso diversi modelli e metodologie computazionali.

c. Il meccanismo di inibizione di un derivato idrolizzato della piperlongumina, noto agente anticancro, nei confronti dell'enzima glutatione transferasi classe pi è stato investigato seguendo diversi cammini di reazione.
d. Un'altra indagine avviata ed in corso di approfondimento è quella inerente ad una proteina ingegnerizzata "de novo protein" , un eptamero *coiled-coil,* (CC-Hept-Cys18-His22-Glu25), svolgente attività esterasica ad opera della triade catalitica Cys-His-Glu usualmente presente in enzimi naturali.

L'attività di ricerca ha consentito la pubblicazione o la stesura dei seguenti lavori:

1 M. Prejanò, T. Marino, N. Russo.

How Can Methanol Dehydrogenase from Methylacidiphilum fumariolicum Work with the Alien CeIII Ion in the Active Center? A Theoretical Study.

Chemistry, A Euopean Journal 2017, 23, 8652-8657. doi: 10.1002/chem.201700381.

2 M. Prejanò, T. Marino, C. Rizzuto, J. C. Madrid Madrid, N. Russo, M. Toscano.

Reaction Mechanism of Low-Spin Iron (III)-and Cobalt (III)-Containing Nitrile Hydratases: A Quantum Mechanics Investigation.

Inorganic Chemistry 2017, 56, 13390-13400. doi:10.1021/acs.inorgchem.7b02121

3 M. Prejanò, T. Marino, N. Russo.

QM cluster or QM/MM in computational enzymology: the test case of LigW-decarboxylase.

Frontiers in Chemistry 2018, 6, 249. doi: 10.3389/fchem.2018.00249.

4. M. Prejanò, T. Marino, N. Russo.

On the inhibition mechanism of glutathione transferase P1 by piperlongumine. Insight from theory.

Frontiers in Chemistry 2018. *Accepted*

5. L. Sgrizzi, M. Prejanò, I. Romeo, T. Marino, N. Russo.

Why nature prefers hydroxy-proline in the deacetylation process promoted by peptide glycan N-deacetylase: insight from molecular simulations.

Submitted to Chemistry, A European Journal.

6. M. Prejanò, F. E. Medina, T. Marino, P. A. Fernandes, M. J. Ramos, N. Russo.

How the destabilization of a reaction intermediate affects enzymatic catalysis: the case of human transketolase.

Manuscript in preparation

7. M. Prejanò, I. Romeo, T. Marino, N. Russo.

Mechanistic insights of hydrolytic activity into a de novo Functional protein framework. *Manuscript in preparation*

Il dott. Prejanò nella sua attività come dottorando, ha dimostrato di possedere una buona capacità ad apprendere metodologie teoriche complesse nel campo della Chimica Teorica e Computazionale. Il dottorando ha seguito i corsi organizzati nell'ambito della Scuola di Dottorato e ha partecipato a seminari tenuti da illustri esperti del settore.

Nel periodo trascorso all'estero presso il Dipartimento di Chimica e Biochimica l'Università di Porto in Portogallo nel gruppo di ricerca di Biochimica Teorica e Computazionale, sotto la supervisione della Prof.ssa Maria João Ramos, il Dott. Prejanò ha avuto modo di approfondire la conoscenza dei metodi di modellizzazione teorica utili per i processi oggetto di studio nonché acquisire competenze utili per la sua formazione complessiva.

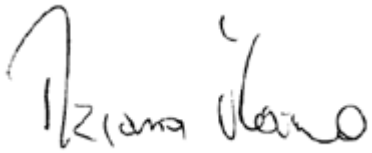
L'attività triennale di ricerca ha prodotto quattro articoli già pubblicati, due articoli in preparazione ed uno sottomesso. Inoltre, l'attività svolta è stata esposta nei seminari semestrali in Italia e in Portogallo, e tramite contributi orali o poster in congressi scientifici nazionali e internazionali.

Nell' affrontare i punti più critici del suo progetto di ricerca e nella razionalizzazione dei risultati, il dottorando ha dimostrato di aver raggiunto un buon grado di indipendenza ed autonomia.

Il Dott. Prejanò nella sua permanenza in laboratorio ha stabilito contatti con tutti i componenti dello stesso e ha dimostrato di possedere una buona attitudine alla collaborazione e disponibilità al confronto ed al trasferimento delle conoscenze apprese.

Il lavoro di ricerca dalla Dott Prejanò è stato svolto e portato a termine con il raggiungimento di gran parte degli obiettivi previsti nel progetto iniziale.

Il giudizio complessivo del Supervisore sulla formazione raggiunta, sull'attività di ricerca svolta nel corso del triennio, sulla personalità scientifica e l'attitudine alla ricerca della Dott. Prejanò è, pertanto, eccellente.

Il Supervisore giudica quindi Mario Prejanò meritevole di aspirare al titolo di Dottore di Ricerca in Medicina Traslazionale, indirizzo Progettazione Molecolare.

Il supervisore
Prof.ssa Tiziana Marino